Transition Metal Catalyzed Water Oxidation: Computational Insight into the Mechanistic Proposals

---

Dissertation

zur

Erlangung der naturwissenschaftlichen Doktorwürde

(Dr. sc. nat.)

vorgelegt der

Mathematisch-naturwissenschaftlichen Fakultät

der

Universität Zürich

von

Mauro Luca Schilling

von

Horgen ZH

Promotionskommission

Prof. Dr. Sandra Luber (Vorsitz)

Prof. Dr. Greta R. Patzke

Dr. Peter Deglmann

Zürich, 2020

Abstract

Fighting man made climate change is one of the great challenges of the 21th century. In this process, the development of renewable energy sources plays a key role. The search for viable solutions led scientists back to Nature, where they got inspired by the light dependent reactions of photosynthesis that uses solar energy to build up and maintain a proton gradient as well as reduction equivalents. The underlying chemical reaction of this process is called 'water oxidation', this is the oxidative half-reaction of the water splitting reaction.

There are numerous compounds, of both homogeneous and heterogeneous nature, known to catalyze said reaction. Unfortunately, the efficiency of most of them is far from optimal, which effectively prevents their application on an industrial scale. To facilitate rational design of more efficient catalysts, an understanding of the underlying reaction mechanism is of utter importance. In the course of this work, homogeneous catalytic systems belonging to two different families of water oxidation catalysts (WOCs) will be introduced. On the one hand, there is the family of bio-inspired Co^{II} based cubanes, a series of highly flexible polynuclear metal clusters that haven been shown to be potent WOCs. On the other hand, there are mononuclear Ru-based catalysts, the polypyridyl ligand framework which has been shown to be an ideal playground for chemists to further improve the catalysts efficiency.

In this work, we apply a broad spectrum of computational methodologies and simulations to further increase the mechanistic understanding of those catalysts. For this purpose, we rely on density functional theory to obtain accurate structural models and properties of possible catalytic intermediates. This not only leads to a more elaborate understanding of the catalytic cycles but also allows us to suggest chemical modifications to existing catalysts in an attempt of rational catalyst design.

With our studies we advance the field of computational catalysis in the context of water oxidation by applying forefront *ab initio* molecular dynamics simulation approaches to account for the dynamics of both the solute and the solvent. This allowed us to gain insights on both the catalysts, the surrounding solvent as well as their crucial interplay at various states of the catalytic cycle. We in particular show how the structural flexibility of the transition metal based catalysts and their non-innocent ligands are influenced by the presence of solvent molecules. By employing enhanced sampling techniques we went from solely describing catalytic intermediates and their surrounding to modeling chemical reactions such as the $O-O$ bond formation in solution. In doing so, we were able to describe the water nucleophilic attack mechanism on metal-oxo species and its associated intermediates at an unprecedented level of detail.

Kurzzusammenfassung

Eine der grössten Herausforderungen, welcher wir uns im 21. Jahrhundert stellen müssen, ist die Bekämpfung des durch den Menschen verursachten Klimawandels. Eine Schlüsselrolle spielt dabei die Entwicklung neuer Technologien im Bereich der erneuerbaren Energien.

Um die damit verbundenen Herausforderungen zu meistern, haben sich Wissenschaftler von der Natur inspirieren lassen. Hierbei ist insbesondere die Lichtreaktion der Photosynthese von Interesse, in welcher die Energie der Sonne benutzt wird, um einen Protonengradienten aufzubauen und Reduktionsäquivalente zu generieren. Dabei handelt es sich um die so genannte Wasseroxidation, die oxidative Halbreaktion der Wasserspaltung.

Bereits heute ist eine grosse Anzahl mikro- und makromolekularer (*homogeneous/heterogeneous*) Katalysatoren bekannt, welche zur Wasseroxidation fähig sind. Zum jetzigen Zeitpunkt sind diese jedoch noch nicht effizient genug, um im grossindustriellen Massstab eingesetzt werden zu können. Es liegt nun an der Wissenschaft den Reaktionsmechanismus dieser Katalysatoren im Detail zu untersuchen. Ziel ist es, die Effizienz bestehender Katalysatoren zu optimieren und die Entwicklung neuer Katalysatoren voranzutreiben.

In dieser Arbeit werden zwei Familien von mikromolekularen Wasseroxidationskatalysatoren vorgestellt. Dabei handelt es sich einerseits um cobaltbasierte, polynukleare Verbindungen, welche dem an der Photosynthese beteiligten Mangankomplex angelehnt sind. Andererseits untersuchen wir mononukleare Rutheniumkomplexe, deren Polypyridin-Liganden seit geraumer Zeit Ziel von chemischen Modifikationen sind, um die katalytische Effizienz zu steigern.

Um den katalytischen Reaktionsmechanismus besser zu verstehen, haben wir eine Reihe von Methoden und Simulationsprotokollen aus dem Bereich der computergestützten Chemie angewendet. Strukturen von möglichen Intermediaten des katalytischen Zyklus und deren Eigenschaften wurden dabei mit Hilfe der Dichtefunktionalstheorie vorhersagt. Dies hat es uns nicht nur erlaubt, ein besseres Verständnis der Eigenschaften der Katalysatoren zu erhalten, sondern wir konnten auch Vorschläge machen, wie deren Effizienz verbessert werden könnte.

In unseren Studien haben wir die Grenzen der computergestützten Katalyse ausgelotet. Wir benutzen *ab initio* Molekulardynamik Simulationen um die Beweglichkeit sowohl des Lösungsmittels als auch des in ihm gelösten Stoffes zu beschreiben. So konnten wir die Interaktionen des Lösungsmittels mit den ver-

schiedenen Intermediaten des katalytischen Zyklus besser verstehen. Dabei haben wir beobachtet wie die strukturelle Flexibilität des Übergangsmetall-Katalysators und seiner Liganden vom Lösungsmittel beeinflusst wird.

Mit Hilfe gerichteter Erkundung der potentiellen Energielandschaft (*enhanced sampling*) konnten wir nicht nur stabile Intermediate des katalytischen Zyklus beschreiben, sondern auch chemische Reaktionen wie die O−O Bindungsbildung in Lösung untersuchen. So haben wir den nukleophilen Angriff eines Wassermoleküls auf den Oxo-Liganden am Metallzentrum im Detail aufklären können.

Disclaimer

In accordance with paragraph 11 article 1 and 2 of the '*Verordnung über die Promotion an der Mathematatisch-naturwissenschaftlichen Fakultät der Universität Zürich (Promotionsverordnung)*' (415.463), most parts of this dissertation have already been published in peer-reviewed journals. Those are listed in the section 'Academic Record', where for each publication the contributions of the author (MS) are given. Further, if the content of the publications is either word-for-word or as a summary re-used in this thesis, a reference to the corresponding section is given. Additionally, at the beginning of each chapter the reference to the corresponding article, where MS is always listed as the first author is given. In general, those chapters are the author accepted manuscripts with an shorted introduction and method section to avoid repetition. Chapter 1 is an extensively modified version of a published book chapter with several additional paragraphs not included in the original manuscript. To avoid content repetitions the article on which Chapter 3 is based was stripped of some of its content and some passages were reformulated.

Academic Record

Articles

- Liu, H.; Schilling, M.; Yulikov, M.; Luber, S.; Patzke, G. R.; Homogeneous Photochemical Water Oxidation with Cobalt Chloride in Acidic Media *ACS Catal.* **2015**, *5*, 4994–4999.

  Contribution: MS participated in the scientific discussions, performed all DFT simulations and contributed to the manuscript preparations.

- Scherrer, D.; Schilling, M.; Luber, S.; Fox, T.; Spingler, B.; Alberto, R.; Richmond, C.J.; Ruthenium water oxidation catalysts containing the non-planar tetradentate ligand biisoquinoline dicarboxylic acid (biqaH_2) *Dalton Trans.* **2016**, *45*, 19361–19367.

  Contribution: MS participated in the scientific discussions, performed all DFT simulations and contributed to the manuscript preparations.
  Declaration of content: A brief summary of the published article is included in Chapter 1.

- Schilling, M.; Patzke, Greta R.; Hutter, J.; Luber, S.; Computational Investigation and Design of Cobalt Aqua Complexes for Homogeneous Water Oxidation *J Phys. Chem. C* **2016**, *120*, 7966–7975.

  Contribution: MS performed the study under direct supervision of SL, performed all DFT simulations and wrote the manuscript.

- Schilling, M.; Hodel, F. H.; and Luber, S.; Discovery of open cubane-core Structures for biomimetic $\{LnCo_3(OR)_4\}$ Water Oxidation Catalysts *ChemSusChem* **2017** *10*, 4561–4569.

  Contribution: MS performed the study under direct supervision of SL based on an existing project by FH. DFT simulations were performed by MS based on structures generated by FH. The manuscript was written by MS.
  Declaration of content: The published article corresponds to Chapter 4.

- Gil-Sepulcre, M.; Böhler, M.; Schilling, M.; Bozoglian, F.; Bachmann, C.; Scherrer, D.; Fox, T.; Spingler, B.; Gimbert-Suriñach, C.; Alberto, R.; Bofill, R.; Sala, X.; Luber, S.; Richmond, C. J. and Llobet, A.; Ruthenium Water Oxidation Catalysts based on Pentapyridyl Ligands *ChemSusChem* **2017**, *10*, 4517–4525.

  Contribution: MS was involved in the supervision of MB who performed the DFT simulations.
  Declaration of content: A brief summary of the article is included in Chapter 1.

- Song, F.; Moré, R.; Schilling, M.; Smolentsev, G.; Azzaroli, N.; Fox, T.; Luber, S.; Patzke, G. R.; $\{Co_4O_4\}$ and $\{Co_xNi_{4-x}O_4\}$ Cubane Water Oxidation Catalysts as Surface Cut-Outs of Cobalt Oxides *J. Am. Chem. Soc.* **2017**, *47*, 14198–14208.

  Contribution: MS participated in the scientific discussions, performed all DFT simulations and contributed to the manuscript preparations.
  Declaration of content: A brief summary of the published article is included in Chapter 1.

- Schilling, M.; Böhler, M.; Luber, S.; Towards the rational design of the Py5-ligand framework for ruthenium-based water oxidation catalysts *Dalton Trans.* **2018**, *47*, 10480–10490.

  Contribution: MB and MS performed the study together under the supervision of SL (shared first authorship). DFT simulations were performed by MB and complemented by MS. The manuscript was written by MS based on the master thesis of MB.[1]
  Declaration of content: The published article corresponds to Chapter 5.

- Schilling, M.; Luber S.; Determination of pKa Values via ab initio Molecular Dynamics and its Application to Transition Metal-Based Water Oxidation Catalysts *Inorganics* **2019**, *7*, 73.

  Contribution: MS performed the study under direct supervision of SL, performed all DFT simulations and wrote the manuscript. Declaration of content: The published article corresponds to Chapter 6.

- Song, F.; Al-Ameed, K.; Schilling, M.; Fox, T.; Luber, S.; Patzke, G. R.; Mechanistically Driven Control over Cubane Oxo Cluster Catalysts *J. Am. Chem. Soc.* **2019**, *141*, 8846–8857.

  Contribution: MS participated in the scientific discussions and assisted KA who conducted the DFT simulations. Declaration of content: A very brief summary of the article is included in Chapter 1.

- Grau, S.; Schilling, M.; Moonshiram, D.; Benet-Buchholz J.; Luber, S.; Llobet, A.; Gimbert-Suriñach, C.; Electrochemically and Photochemically Induced Hydrogen Evolution Catalysis with Co-Tetraazamacrocycles Occur via Different Pathways *submitted*.

  Contribution: MS participated in the scientific discussions evolving around the project as well as performed all the simulations.

- Schilling, M.; Cunha, R. A.; Luber, S.; Zooming in on the O−O Bond Formation - an *ab initio* Molecular Dynamics Study Applying Enhanced Sampling Techniques *submitted*.

  Contribution: MS conducted the study under the supervision of SL and with scientific and methodical inputs of RC. MS also preformed all the DFT simulations and wrote the manuscript.
  Declaration of content: The manuscript corresponds to Chapter 7.

Review Article

- Schilling, M.; Luber, S.; Computational Modeling of Cobalt-Based Water Oxidation: Current Status and Future Challenges *Front Chem* **2018**, *6*, 100.

  Contribution: Contribution: MS wrote the manuscript under supervision of SL.
  Declaration of content: The published article corresponds in large parts to Chapter 3.

Book Chapter

- Schilling, M.; Luber, S.; Insights into artificial water oxidation – A computational perspective *Adv. Inorg. Chem.* **2019**, *74*, 61-114.

  Contribution: MS wrote the manuscript under supervision of SL.
  Declaration of content: The published article corresponds to Chapter 1.

Acknowledgment

First and foremost, I would like to state by gratitude to my supervisor Prof. Dr. Sandra Luber. She encouraged me to peruse a PhD under her supervision after our worthwhile collaboration during my master thesis. Together we continued our scientific journey in the field of artificial water oxidation where Sandra always know how to steer my broad interest to challenging topics. Further I would like to thank Prof. Dr. Greta R. Patzke for the many fruitful collaborations over the years. I am also grateful for the first year of my PhD that was hosted by Prof. Dr. Jürg Hutter, before Sandra became independent as a SNF professor. For the administration of the PhD program and the organization of many great scientific and networking events I would like to thank the 'Graduate School of Chemical and Molecular Science Zurich' (CMSZH).

My special gratitude goes to 'my' master student Michael with whom I enjoyed working and arguing at eye level. I would also like to thank the many great people with whom I had the pleasure to work or to spent time with over the past couple of years. Starting with the current and past Luber Group members: Edward, Eva, Johann, Katharina, Karrar, Mauricio, Momir, Rahele, Ruocheng, and Tomaš with whom I had and still have a great time. Then the member of the Hutter Group, especially: Augustin, Florian, Jinggang, Michela, Patrick and Tiziano, as well as the members of the Sigel Group: Fabio, Richard B. Richard C., Silke and Sofia with whom I enjoyed many great discussions during lunch and coffee breaks. Finally, I want to acknowledge the members of the Patzke Group: Fangyuan, Jingguo, and Karla for many interessting dialogues.

I also would like to thank my various collaborators especially Craig Richmond and Carolina Gimbert-Suriñach for the challenging questions they have raised. Finding answers to them often resulted in many new ideas some of which ended up in further projects. Most of my work would not have been possible without the efforts of the IT-Support provided by Andreas Glöss and Sascha Giger.

Last but not least I would like to thank my family, my 'attachment' and of course my friends, among them I want point out my various climbing buddies. All the aforementioned people constantly supported me and encouraged me to push forward in all aspects of life.

CMSZH
CHEMICAL MOLECULAR SCIENCES ZURICH
GRADUATE SCHOOL

Mauro Schilling
December 2019

Contents

Acronyms

AIMD *ab initio* molecular dynamics.

AP associated product.

AR associated reactant.

bda 2,2′-bipyridine-6,6′-dicarboxylate.

bipy bipyridine.

biqa [1,1′-biisoquinoline]-3,3′-dicarboxylate.

BO Born-Oppenheimer.

CAN cerium-ammonium-nitrate.

CASSCF complete active space self-consistent-field.

CCSD(T) coupled-cluster singlet doublet (triplet).

CI climbing-image.

CI configuration-interaction.

CN coordination number.

COSMO conductor-like screening model.

COSMO-RS conductor-like screening model for realistic solvents.

CPMD Car-Parinello molecular dynamics.

CSM continuum solvent model.

CV collective variable.

D3 Grimmes D3 empirical dispersion correction.

DCOSMO-RS direct-COSMO-RS.

DFT density functional theory.

dpyC(OH)(O) hydroxyldi-2-pyridinylmethanolato.

ET electron transfer.

EVB empiric valence bond.

FES free energy surface.

GASSCF generalized active space self-consistent-field.

GGA generalized gradient approximation.

HF Hartree-Fock.

hmp 2-(hydroxymethyl)-pyridine.

HOMO highest occupied molecular orbital.

I2M interaction of two metal oxo.

isoq 6-bromoisoquinoline resp. isoquinoline.

KS Kohn-Sham.

LDA local density approximation.

LUMO lowest unoccupied molecular orbital.

MCSCF multi-configurational self-consistent field.

MD molecular dynamics.

MEP minimum energy pathway.

MM molecular mechanics.

NEB nudged elastic band.

O2DI oxygen dissociation.

OAc acetate.

OEC oxygen-evolving-complex.

P product.

PCET proton coupled electron transfer.

PCM polarizable continuum model.

pda phenantroline-2,9-dicarboxylate.

pic picoline.

PMF potential of mean force.

POM polyoxometalate.

PT proton transfer.

py pyridine.

Py5R 6,6″-(R(pyridin-2-yl)methylene)-di-2,2′-bipyridine, where R is either a methoxy- or methyl group.

pyMeO pyridin-2-ylmethanolato.

QM quantum mechanics.

R reactant.

RASSCF restricted active space self-consistent-field.

RC radical coupling.

RDF radial distribution function.

RESP restricted electrostatic potential.

SCF self-consistent field.

SHE standard hydrogen electrode.

TOF turn-over-frequency.

TON turn-over-number.

TS transition state.

TS-vdW Tkatchenko and Scheffler van der Waals method.

VV Vydrow Van Voorhis non-local correlation functional.

WNA water nucleophilic attack.

WOC water oxidation catalyst.

XC exchange-correlation.

ZPE zero-point-energy.

Chapter 1

Theoretical insights into artificial Water Oxidation

This chapter is modified version of: Schilling, M.; Luber, S.; *Adv. Inorg. Chem.* **2019**, *74*, 61-114.[2] If not stated otherwise all figures of this chapter were published in the reference mentioned beforehand.

1.1 Introduction

1.1.1 Artificial Water Splitting - a Short Overview

Almost on a daily basis, the news report extreme weather phenomena or new records in terms of temperature, rainfall, or winds. Some of the extreme conditions, such as droughts, are believed to be the direct or indirect consequences of man-made climate change.[3–5] In 2015, most countries in the world agreed to take actions against man-made climate change by signing the *Paris agreement*, which has the aim to limit the rise of the average global temperature to below 2°C above pre-industrial era. Especially, by reducing the emission of greenhouse gases as well as by the development of clean, renewable energy sources.[6] Among others, photovoltaic modules are used to harvest energy provided by the sun and convert it to an electric potential. However, the efficiency of the those devices is strongly dependent on the latitude, the local climate (e.g. seasons), the elevation and of course the daytime. In particular, the daily peak in energy production around noon causes some problems, since the energy consumption of mankind normally peaks in the evening. This results in a mismatch in production and consumption, which is often referred to as 'Duck curve'.[7] In order to overcome this problem, development of energy storage and transport capabilities have seen a lot of scientific interest in the recent past.

Nature has already solved the problem of energy storage billions of years ago with a process called photosynthesis. This process describes a combination of photo- and electrochemical reactions performed by a complex array of enzymes that - in an ingenious assembly - take advantage of compartmentalization within cells. Chemically, photosynthesis can be divided into two parts - water oxidation and CO_2 reduction. The first part, water oxidation, is related to conversion of sunlight energy into a chemical potential, which is 'stored' in the form of chemical bonds in the second part. An analogous process - water splitting - is artificially achieved by electrolysis where water is split into O_2 and H_2. The efficiency of this process relies heavily on the nature of the electrode, which is often composed of platinum. The scarce

availability of rare metals limits economically viable implementation of this process on a global scale. Further, in its classical implementation electrolysis is driven by an electric potential rather than by solar energy. The use of electric energy produced from renewable sources is possible, however this is an indirect pathway which may contain several interfaces between different devices. Complex assemblies of multiple devices are in general prone to losses. Those losses can sum up to an overall efficiency which may not be economically sustainable yet. It would therefore be desirable to develop processes where sunlight is directly stored in chemical bonds - such processes are called 'artificial photosynthesis'. This term is somewhat ambiguous, since it does not specify in which form i.e. in what kind of chemical bonds the sunlight energy is stored, be it H_2, or as in nature in the form of C–H bonds.

In the field of heterogeneous catalysis there are already systems known which are up to the task. Some of them such as Nocera's wireless solar water splitting device also known as 'artificial leaf' are promising to do a decent job.[8,9] However, the main drawback of such complex assemblies is that a systematic improvement of specific properties like the catalytic efficiency or lifetime is often very difficult. Therefore researchers often focus only on a single part of the whole process, be it light harvesting, water oxidation or water reduction, in the hope to master each piece individually and combine the optimized solutions in the end. However, the heterogeneous nature of the catalysts used to drive those processes often does not allow for a mechanistic insight on a molecular level, knowledge of which might be crucial in order to understand the advantages and disadvantages of the use of certain materials, and could in principle help to systematically improve existing catalysts or even serve to develop guidelines for the design of new catalysts.

Because of these limitations homogeneous catalysts might be a suitable alternative. Their well defined molecular structure in principle allows for in-depth mechanistic studies that could lead to the identification of potential bottlenecks in the catalytic process. Of course homogeneous catalysts also come with a drawback - many of them suffer from stability issues under catalytic conditions, and addressing these is often challenging. As for the heterogeneous catalysts the final goal is their integration into a single device. Attempts in this direction have already been made, among them is a functioning water-splitting device based on an immobilized, but not covalently linked Ru based water oxidation catalyst (WOC) by Sun and co-workers.[10] However it is a mere proof of concept rather than a scalable device which may soon enter application on an industrial scale. For a an overview of the current state of the art devices and on immobilization techniques see Yang *et al.*[11]

1.1.2 Water Oxidation Reaction

In the following the attention will be focused on the oxidative part of the water splitting process which is often thought of to be the bottleneck of the whole process. The overall chemical reaction for water oxidation is:

$$2\,H_2O \longrightarrow 4\,H^+ + O_2 + 4\,e^-. \tag{1.1}$$

The half-reaction itself is endothermic, which means that 113.5 kcal mol^{-1} (4.92 eV, at standard conditions) in terms of free energy are required in order to drive the reaction. This is the thermodynamic minimum, assuming there are no further kinetic

barriers for chemical reactions such as the O−O bond formation, reactant association or product release, which can only be achieved in a hypothetical ideal system. In practice those barriers turned out to be significant, in particular due to the complexity of the reaction which consists of four consecutive oxidation and deprotonation reactions and an O−O bond formation. To match the thermodynamics of an ideal process is on a purely theoretical level the main goal in catalyst development. However, a catalyst that has the potential to play a crucial role in upcoming energy supply chains based on renewable energy sources has to fulfill other requirements as well. Those might be defined as:[12,13] (1) efficiency - high activity under mild conditions (neutral pH, room temperature, atmospheric pressure); (2) specificity - towards a single product which might be easy to separate from the reaction mixture; (3) durability - longterm stability in order to achieve high turnover numbers or a facile recycling and regeneration process; (4) cleanliness - environmentally friendly i.e. 'green chemistry', non-toxic components; (5) adaptability - easy to modify e.g. for different conditions, to immobilize on surfaces, or to couple with different catalysts. Most of those criteria are interdependent and cannot be individually targeted when developing novel catalysts. Therefore in a first step, one often primarily aims for efficiency. Here, theoretical studies can be crucial in order to understand the fundamental properties of the catalyst, in particular, when attempting to lower the activation barrier of the rate-determining step. This requires knowledge on intermediates that are part of the catalytic cycle. The elusive nature of those intermediates often makes theoretical modeling a precondition of any conclusive mechanistic proposal.

In the past decade numerous molecular WOCs have been proposed, most of them containing transition metals - primarily Mn, Fe, Co, Ru, and Ir. However there are also examples known that contain Ni or Cu.[14–19] For some of those catalysts there exist mechanistic proposals which are either based on experiments, theory or in the best case a combination of both.[20–25]

While the detailed mechanisms are dependent on the nature of the catalyst, all of them share some common features. We reviewed the mechanistic studies on molecular cobalt based WOCs, where we highlighted the similarities of the different catalysts in terms of their proposed mechanisms (see Chapter 3).[13] In order to lay the foundation for the following sections, we will give a brief overview of the most general concepts in terms of water oxidation and on how to model this process.

1.2 Water Oxidation - a Computational Perspective

Catalytic water oxidation can be divided into the following stages: (1) water activation, i.e. oxidation and deprotonation in order to reach a metal-oxo (M=O) intermediate, (2) oxygen-oxygen bond formation, for which there are two general possibilities: a water nucleophilic attack (WNA) or an interaction of two metal-oxo species (I2M), sometimes also referred to as radical coupling (RC); depending on the nature of the catalysts there are many variations of those two reactions. (3) release of O_2, determined by the O−O bond formation mechanism; either further oxidations are required or O_2 is directly released from the catalyst. (4) regeneration of the catalyst by associating a water molecule. This step might happen in a concerted fashion with the previous step. A generalized catalytic cycle is shown in Figure 1.1.

The initial stage of the catalytic cycle is characterized by electron transfer (ET) and

proton transfer (PT) reactions or *vice versa*. In order to avoid the accumulation of charge, ET reactions are normally directly followed by PT reactions. If those two happen together in a concerted process then they are referred to as a proton coupled electron transfer (PCET). Several ETs and PTs or PCETs lead to a metal-oxo species, which is a prerequisite for the O–O bond formation. The origin of the high reactivity of those metal-oxo species is often uncertain as illustrated by the following example.

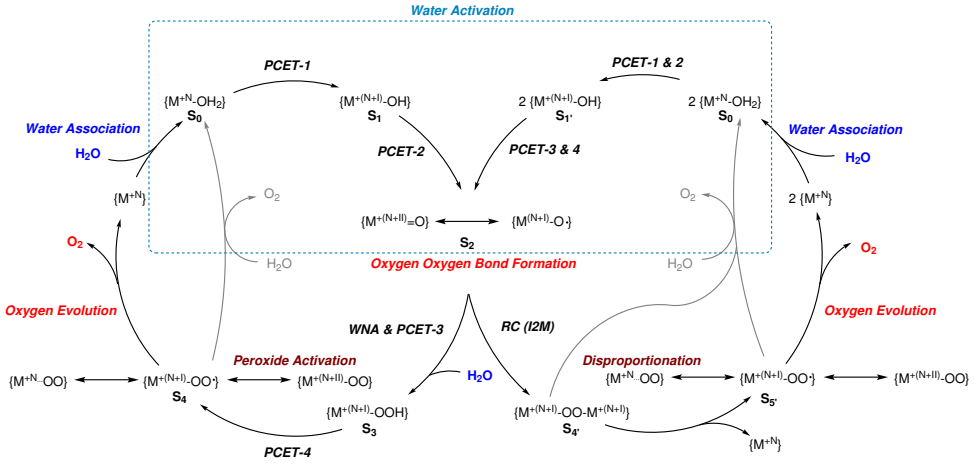


Figure 1.1: Schematic overview of a possible water oxidation mechanism. M represents any redox active transition metal, N as a superscript represents the formal oxidation state of the metal center in the catalytic resting state. Reactions connecting two states do not necessarily have to be PCETs, various sequences of ET and PT reactions are possible. Further, some catalysts might require an activation step to reach the catalytic ground state.

For the sake of simplicity, let us assume that a model catalyst is only composed of a single metal center, which in its resting state bears a single aqua ligand - $\{M{-}OH_2\}^{n+}$. After two consecutive PCETs, the metal-oxo species $\{M{=}O\}^{n+}$ is reached. The Lewis formula of this species is somewhat misleading, depending on the electron counting method - ionic ($M^{+(n+II)} + O^{-II}$) or neutral ($M^{+n} + O^{0}$) - a different formal oxidation state might be assigned to the metal center. Further, for the metal-oxo species there are resonance structures i.e. $\{M{=}O\} \longleftrightarrow \{M{-}O^{\bullet}\}$, suggesting another possible formal oxidation state of $M^{+(n+I)} + O^{-I}$. The oxyl formalism implies on the one hand a weaker metal-ligand interaction, and on the other hand a higher reactivity of the ligand due to its radical character. The metal-ligand interaction is expected to increase for highly oxidized species i.e. n = 5 or higher, where most of the metal's 3d orbitals are empty. For those species one might draw the structure as $\{M{=}O\} \longleftrightarrow \{M{\equiv}O\}$, where the formal oxidation states are $M^{+(n-I)} + O^{+I}$. Here, the +I oxidation state of the ligand highlights its electrophilic character, in the case of neutral counting.[26]

We have to keep in mind that oxidation states are a mere bookkeeping tool for the number of electrons in the system. There is no rigorous quantum chemical definition of oxidation states and in particular in transition metal complexes assignment of formal oxidation states might lead to contradictory interpretations. However,

theoretical chemistry offers several techniques to deduce formal oxidation states from the total electron density or wavefunction such as Bader, Voronoi, Mulliken, or Löwdin charges to mention the most popular ones (see Sit *et al.* and references therein).[27] More advanced concepts are based on the localization of the electron density onto various kinds of atom centered orbitals.[27–31]

1.2.1 Mechanistic Proposals for the Oxygen-Oxygen Bond Formation

The WNA and RC mechanisms were already briefly introduced in the previous sections. In this paragraph the focus is laid on the different variations of those mechanisms (see Figure 1.2).[32] While in principle there are several reaction pathways imaginable for each catalyst, usually only a few of them are plausible under the applied reaction conditions. Ideally water oxidation takes place at a neutral pH, however depending on the catalyst and / or the use of photosensitizer or sacrificial oxidants, different conditions might be used. For example, the performance of most Ru-based WOCs is evaluated in the presence of the chemical oxidant cerium-ammonium-nitrate (CAN) under acidic conditions (pH = 1). In the case of an intermolecular WNA, a water molecule from the solvent has to attack nucleophilically the metal-oxo species, forming a formal $\{M-OOH_2\}$ species, which presumably quickly deprotonates to form the corresponding $\{M-OOH\}$ species. However, it is probably more realistic to assume that the oxygen-oxygen bond formation and the deprotonation take place in a concerted manner. Further, the weak nucleophilicity of the water molecules might be enhanced by the presence of a base, be it intra or intermolecular. In this process the base does not necessarily have to interact with the nucleophile itself. An indirect participation in a proton reshuffling process during the O−O bond formation might also be possible, as has been suggested for several catalysts.[33–37]

When the intramolecular base is not an integral part of the ligand framework, but rather a species such as a hydroxide coordinated to a second metal center, as it might be the case in heterogeneous or polynuclear catalysts, then more options for the O−O bond formation arise. First, the hydroxido ligand might act as a nucleophile in an intramolecular WNA. Secondly upon further oxidation a second metal-oxo might be formed which could undergo the O−O bond formation in an intramolecular RC fashion. An RC mechanism is also possible if the O−O reaction is bimolecular, i.e. when two $\{M=O\}$ species dimerize to form the O−O bond. The ability of catalysts to dimerize is strongly dependent on their respective ligand spheres, that should be able to stabilize a face-to-face orientation of the $\{M=O\}$ intermediates.

Experimentally it is difficult to distinguish between the different reaction pathways discussed above. Often the only easily accessible information is the reaction order derived from the rate equation. This, however only allows to differentiate a second-order reaction such as a O−O bond formation in an intermolecular RC from all the other possibilities shown in Figure 1.2 as all of them are first-order in terms of the catalyst concentration. Another common experiment is the use of isotopically labeled water during the synthesis of the catalyst or as a solvent during water oxidation, in order to identify the source of the released O_2 .[38] However, the extent to which such studies can help to identify the O−O mechanism strongly depends on the system at hand.

While dangling pyridyl subunits are commonly expected to act as an intramolec-

ular base, there are also other mechanistic proposals in literature. For exampling, Pushkar *et al.* proposed that a pyridyl-subunit could be oxidized to the corresponding pyridyl-N-oxide. This species could potentially act as the nucleophile in the O–O bond formation by RC.[39] The validity of this mechanism has still to be proven, however the fact that the ligand framework of most WOCs contain pyridyl subunits makes the mechanism appealing. There are already the first theoretical studies appearing where the oxygen-atom-transfer from the metal to the pyridine has been modeled.[40]

A fundamentally different pathway for the O–O bond formation was proposed by Zhan *et al.* They proposed a nucleophilic attack of a hydroxide onto the carboxylate forming a carbonate like intermediate that transfers one of its oxygen atoms to the metal oxo-species forming the O–O bond.[41] While experimentally difficult to prove, this mechanistic proposal highlights the possibility of the ligands to actively participate in the catalytic cycle i.e to be non-innocent, a fact that is often overlooked.

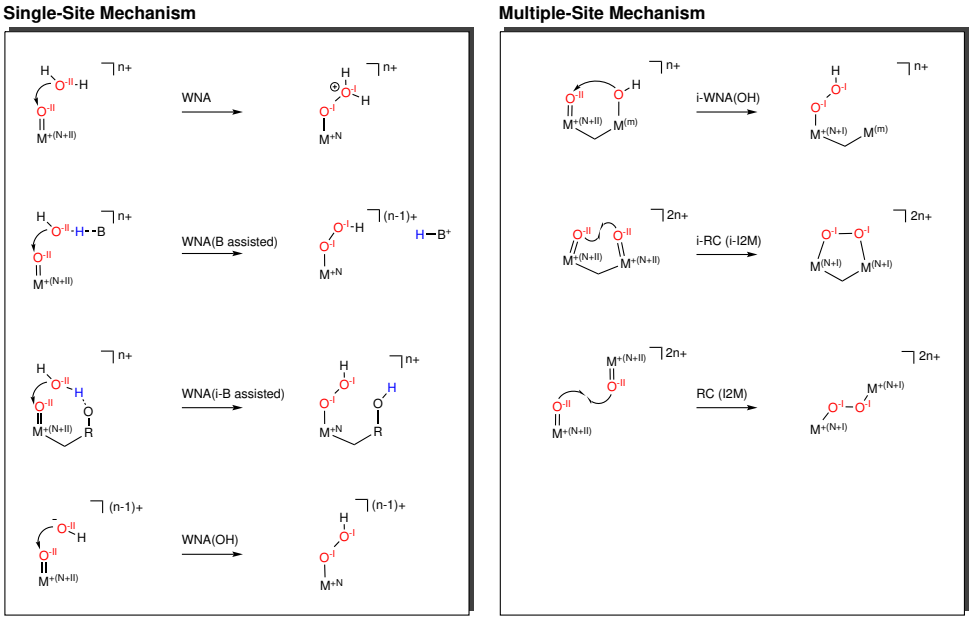


Figure 1.2: O–O bond formation mechanisms, where M represents any redox-active metal center in the formal oxidation state indicated by the superscript, where +N is used for the 'active' metal center and +m for any spectator metal center. Further the overall charge of the species is denoted as n+ and 2n+ in case of a bimolecular reaction. The formal oxidation state N and the overall charge n represent always the same numerical value. The overall charge of the association complex as well as of the product are given to highlight charge conservation. For the sake of clarity the formal oxidation state of the involved oxygen atoms is given in both the reactant and product state. B is a general base in solution, and R is any residue including other transition metals. Reproduced from Schilling and Luber.[13]

1.2.2 Energetics of the Oxygen-Oxygen Bond Formation

Theoretical studies can provide much-needed insight by comparing different routes in terms of their energetics. Besides the thermodynamics, the kinetics of chemical

steps are also of interest, knowledge of which can be gained from transition states (TSs) connecting possible intermediates of the catalytic cycle. Calculating TSs is a non-trivial task and often requires chemical intuition. Nevertheless, there are several standard protocols available. In the following the basic principles of some selected methods will be briefly introduced.

TSs are saddle points on the potential energy hypersurface. The most common approach to converge a structure to such a point is to calculate the second derivatives (Hessian matrix), identify an eigenvector with a negative eigenvalue that corresponds to the reaction coordinate and maximize the energy with respect to this eigenvector, while minimizing it in all other dimensions. This method usually requires a good initial guess that is already close to the TS. If only the reactant (R) and product (P) of a reaction are known, then methods like self-consistent reaction path optimization,[42] or nudged-elastic-band (NEB) methods might be applied.[43,44] In order to find the minimum energy pathway (MEP) connecting two minima, i.e. the R and the P, several intermediate structures (also referred to as beads) along this pathway have to be optimized. In the case of NEB calculations, those beads are connected by springs which ensure an even spacing along the pathway. The energy of those beads is minimized with respect to all degrees of freedom perpendicular to the springs. When the climbing-image (CI) extension to the NEB methodology is used, where the bead with the highest energy is allowed to move along the band, then the converged MEP is guaranteed to contain the saddle point.

Both methods discussed so far rely on the assumption that the reaction coordinate is rather simple and might even be guessed by experienced chemists. This disadvantage might be overcome or at least reduced by using sampling techniques based on molecular dynamics or Monte Carlo. In the limit of infinite sampling time, such sampling techniques would allow to reconstruct the whole potential energy surface. However, in order to limit the computational effort enhanced sampling techniques might be applied that force the systems to explore primarily areas of interest. There are numerous techniques for this purpose, here only one of them - metadynamics (MetaD) will be discussed.[45,46] This technique is based on the continuous addition of an artificial bias potential during a molecular dynamics simulation in order to discourage revisiting the same state again.

States are defined in terms of collective variables (CVs) i.e. components of the reaction coordinate that are able to distinguish unambiguously between the R and P. *A priori* there is no previous knowledge on the reaction path or TS required, however MetaD simulations depend on many parameters, such as the height and width of the Gaussian that are continuously added to the bias potential, the frequency in which those Gaussian are spawned and of course the CVs which define the dimensions of the hypersurface to be explored. Even properly set parameters sometimes do not lead to convergence of the free energy within a reasonable time frame, therefore there are many extensions such as well-tempered and transition-tempered metadynamics (WT-MetaD) that might help to facilitate this process in certain cases.[47–49] For an exemplary application and a brief introduction see Section 2.4.2 and 1.4.4 as well as Chapter 7.

1.2.3 Theoretical Description of Metal-Oxo Species

By drawing Lewis structures one might oversimplify the description of the metal-oxo species too. While $\{M-O^{\bullet}\}^{(n)+}$ implies that there is an unpaired electron, it by no means gives any information on its localization with respect to the oxygen atom. This becomes obvious when it comes to the description of real catalysts such as a triplet $Mn^{V}=O$ complex, that was studied by Ashley and Baik.[26] Based on their theoretical calculations, they assigned the two unpaired electrons that define the triplet state to the non-bonding molecular orbital with d_{xy} symmetry and π^{*}_{M-Oxo} molecular orbital, visualizing of which is beyond the capabilities of Lewis structures.

Therefore the possibility to draw several different resonance structures asks for a solid quantum chemical description in order to pin-down the 'true' nature of the species. Nowadays, the workhorse of quantum chemistry is Kohn-Sham density functional theory (KS-DFT). However, it is a single determinant method, as such DFT usually struggles describing correctly structures like the oxo/oxyl if the real structure is a mixture of these two structures. In order to describe this ambivalence multi-configurational self-consistent field (MCSCF) approaches such as complete active space self-consistent-field (CASSCF) ,[50] restricted active space self-consistent-field (RASSCF),[51,52] and generalized active space self-consistent-field (GASSCF) calculations can be used.[51,53,54] Those methods are non-trivial in their application and are due to their high computational cost only applicable for smaller systems or for very specific questions. It is therefore common to use MCSCF calculations as a benchmark for DFT calculations employing different exchange-correlation functionals. A common observation is that the use of hybrid functionals is required, and that especially the amount of Hartree–Fock (HF) exchange plays a crucial role when it comes to describe metal-oxo species.[26,55] This is also true for activation energies for the O−O bond formation. Ashley and Baik reported for both the WNA- and the RC-mechanism of a Mn-based WOC a non-trivial relationship between the % of HF-exchange and the barrier height. In case of a WNA, they found the activation energy to be a function of the amount of HF-exchange, i.e. larger HF contributions led to a higher activation energy. At the same time, the energetic difference between barriers obtained for high-spin state (in their case triplet) and low-spin state (here singlet) diminished and both spin-states converged to the same barrier. For the RC mechanism, the energetic difference between barriers obtained for HS and LS configurations also decreased but it did not vanish within the probed range.[26] Inspecting the oxyl character of those species revealed that the barrier of neither the WNA nor the RC mechanism was directly dependent on the oxyl character i.e. there was no obvious benefit of a large amount of radical character.

Together with the ever growing computational resources and recent developments such as the domain-based local pair natural orbital framework, even wavefunction based methods such as coupled-cluster singlet doublet (triplet) CCSD(T) become feasible for transition metals, which might in the future be a viable complement to MCSCF methods.[56–58]

1.2.4 How to Account for the Surrounding Solvent?

In general one differentiates between implicit and explicit solvation. Implicit solvation aims to capture the average effects of solute-solvent interactions but has shortcomings when it comes to describe directed interactions such as hydrogen bonding.

While there are different formulations of such solvation models, such as the well known polarizable continuum model (PCM) and conductor like screening model (COSMO),[59] their basic principle is similar. The solvent is approximated as an electrostatic potential interacting with the potential created by the solute. The models differ in the formulation of those interactions as well as the description of the interface region between the charged cavity surrounding the solute and the solvent continuum.[60–63] Further, those solvation models rely on an empirical parametrization to reproduce experimental solvation free energies. The parametrization is often based on a limited set of small to medium sized (neutral) organic molecules and is done at a rather low level of theory, e.g. small basis set.[64,65] Thus the application of implicit solvation models to more complex systems such as transition metal complexes might not in all cases be justified.[66] For a brief introduction of the COSMO methodology see Section 2.2.

While implicit solvation models are known to perform reasonably well for systems with no direct solute-solvent interaction, they often struggle when such interactions become important. This is for example the case for the O−O bond formation by a WNA mechanism. Unless there is a basic site in close proximity to the newly formed O−O bond, a high energy $\{M-OOH_2\}$ species would be formed. It has therefore become customary to incorporate several water molecules into the model systems, that might act as proton acceptor or proton shuttles.[33–35,37,67] The atomistic description of the solute-solvent interactions is referred to as explicit solvation, and might also be combined with implicit solvation to account for the interaction with the bulk solvent. Such a static approach using geometry optimizations is normally limited to a small number of explicit solvent molecules, because solvent molecules do not only interact with the solute but also with other solvent molecules. The various possible solvent configurations can lead to a strongly corrugated energy surface, when comparing energies of different solutes. Hodel *et al.* explored the limits of this approach in some of their studies which are discussed later on in detail (see Section 1.3.2).[68] With COSMO for realistic solvents (COSMO-RS) an attempt was made to overcome the limitations of a 'simple' implicit solvation model. The method is based on quantifying intermolecular electrostatic and hydrogen-bond interactions using the COSMO polarization charge densities, resulting in a more realistic description of the solvent.[59]

Even though sampling techniques such as molecular dynamics come with extraordinarily high computational costs, they are often the method of choice if solvation effects are to be treated explicitly in a dynamic manner. This is true in particular for the O−O bond formation by a WNA where the nucleophile stems from the solvation shell. So far, the immense computational cost and the non trivial application of those methods, in particular in the case when both the catalyst and the solvent are described by DFT, have effectively limited their application to complex systems in the field of homogeneous transition metal-based catalysis. Nevertheless a few examples can be found in literature.[37,68–70]

Originally, sampling techniques such as molecular dynamics were developed in the context of an entirely classical description based on molecular mechanics (MM). However, transition metal complexes that participate in e.g. bond-formation and -breaking reactions, which may go along with a change of the local electronic structure at the metal center, are difficult to describe in a classical sense by means of a force field. Nonetheless, the first classical simulations of a solvated Ru-based WOC

appeared,[41,71,72] where the empiric valence bond (EVB) approach was used to obtain a more detailed picture of the O–O bond formation by an intermolecular RC mechanism.[73]

In order to represent the electronic structure of transition metal complexes sufficiently accurately DFT or post-HF methods are usually required. Both worlds, i.e. quantum mechanics (QM) and MM might be combined in the so called QM/MM methodology, where different parts of the system are calculated at a different level of theory. For example, in the case of solvated transition metal complexes, the solute i.e. the catalyst, with our without the first solvation shell can be described employing DFT while the solvent molecules can be treated classically. The QM/MM methodology performs in particular well when the studied reactions such as the O–O bond formation are intramolecular, i.e. there is no bonding interaction between the solute and solvent. In this case additional insight on the dynamics of the solute might be obtained at significantly reduced cost compared to a full *ab initio* molecular dynamics (AIMD) setup.[74]

1.2.5 Reduction Potentials and pK_a Values

The choice of the solvation model does not only affect the computational results of catalytic mechanisms, related structures and energies, but also other key properties such as reduction potentials and pK_a values. Even though the access to those properties is experimentally often restricted to a limited set of species, theoretical estimates are still of great importance when it comes to judging which species might be present in solution during catalysis.

Depending on the solvation model there are several protocols available to calculate reduction potentials and pK_a values.[75,76] When employing an implicit solvation model, usually thermodynamic cycles, also known as Born–Haber cycles, are used. Those cycles avoid the calculation of a formal H^+ or e^- in solution, by calculating their respective solvation free energies instead.[77] In case the reaction is formulated in an isodesmic manner, either a proton or an electron is transfered to a corresponding base or oxidant instead of being released into the solution. This might lead to more favorable error cancellation, if all the participants of the reaction are similar in their composition, and if they are calculated at the same level of theory.[78] If solvent–solute interactions play a significant role in determining the structure of the reactants and/or the products, then thermodynamic cycles - that rely on the assumption that the chemical structure of a species remains the same upon solution - are not applicable anymore.[79] For certain systems such as transition-metal ions, it turned out that not only the molecules in the first coordination sphere but also those in the second shell have to be described explicitly in order to reproduce experimental reduction potentials.[80] As already discussed, the inclusion of explicit solvent molecules in principle makes the use of sampling techniques necessary. An alternative might be COSMO-RS, which has been shown to be able to reproduce experimental reduction potentials with a better accuracy than other common implicit solvation models, even without inclusion of molecules in the second coordination sphere.[81] A drawback of the methodology discussed above is its evident failure for highly charged species, for which solvent continuum models are usually not parameterized. This limitation can to some extent be overcome if the charged complexes are neutralized for example by the addition or the removal of protons. The reduction potential of the highly charged species, can then be obtained by calculating

several thermodynamic cycles consisting of intermediate species that carry a smaller molecular charge. This methodology was introduced by Srnec and co-workers and is referred to as variable-temperature H-Atom addition/abstraction (VT-HAA).[82]

The 'gold standard' using DFT or wavefunction based methods is still a completely atomistic description of the solvent. Sprik and co-workers developed a DFT-MD and free energy perturbation scheme based protocol that allows the calculation of reduction potentials and pK_a values.[83] A complete recapitulation of this method would be beyond the scope of this work, therefore only the fundamental concept will be introduced. For a more detailed description we refer to the corresponding original literature.[84–88] The scheme has strong ties to Marcus theory,[89] and is conceptually based on the calculation of the free energy of insertion of an electron or proton into the simulation box. This is done by a thermodynamic integration scheme of the vertical energy gap between the two states (with and without an electron or proton respectively). In practice this means that not only the initial state and final state, but also intermediate states have to be sampled, in order to calculate the pK_a or reduction potential of a certain species. This is in particular true for the insertion of a proton where the solvent response is non-linear. Due to the requirement of a reference state within the same methodology and in particular within the same simulation box, a computational hydrogen electrode ($H_2 \longrightarrow 2\,H^+ + 2\,e^-$) has to be calculated as well. In literature this protocol is referred to as insertion–deletion scheme.[90] AIMD simulations are usually performed employing generalized gradient approximation (GGA) exchange-correlation functionals, which are known to perform worse in terms of energy differences than hybrid functionals. However, due to their high computational cost the use of hybrid functionals for MD simulations is currently often out of the question. It is therefore recommended to recalculate the vertical energy gap employing a hybrid functional. This adds even more cost to an already expensive methodology, which in fact limits its applicability to small systems.

In case neither reduction potentials nor pK_a values are of interest but rather dehydrogenation free energies i.e. the free energy difference of a PCET, then certain shortcuts can be taken since there is no need to calculate the computational hydrogen electrode, which makes the protocol more affordable. But why should one be interested in PCETs instead of reduction potentials and pK_a values? Every water oxidation mechanism can be formulated as sequence of four PCETs in order to avoid the accumulation of charge which presumably leads to high energy intermediates. Hodel and Luber successfully applied the above mentioned protocol to calculate dehydrogenation free energies for a model system of a transition-metal based WOC, a solvated Co^{II}-ion, and thereby shown that in the future this methodology might become viable to investigate transition metal catalyzed water oxidation.[91]

Prior to the development of the insertion–deletion scheme a slightly less elaborated protocol, known as *Bluemoon* ensemble, was established by Sprik and Ciccotti.[92–94] The basis of which is to describe the chemical transformation of interest, usually a conformational change or a bond formation/breakage, by a series of constraint MD simulations. Integration of the force profile obtained by fitting the average forces required to maintain the constraint during the individual simulations, results in the free energy of the chemical process. For a brief introduction of the methodology see Section 2.4.1. The *Bluemoon* ensemble has been successfully applied to both chemical reactions such as bond formations and deprotonation reactions,[95] which were

used to estimated p$K_a$values of various compounds.[96,97] There are several proposals how exactly the free energy difference obtained by the *Bluemoon* ensemble might be related to the pK_a. We have compared those proposals by simulating the p$K_a$value of transition metal complexes (see Section 1.4.4 and Chapter 6 for more details).[98]

In the last few paragraphs we shed light on some reoccurring aspects that are important when modeling catalytic systems. We aimed neither to give a full scope of all available methods and protocols nor to give an in-depth description of presented ones.

With this general introduction on transition metal catalyzed water oxidation the foundation for the following sections was laid out, where selected contributions made by Luber and co-workers in the context of water oxidation will be highlighted. The first paragraphs serve as an introduction of the model systems and give a taste of the work done by Hodel and Luber. This is necessary in order to grasp both the complexity of the systems as well as the capabilities of the simulation protocols. In the later paragraphs there is a fluent transition to works with contributions of the author. While in this chapter only a brief summary of the original work by the author are given, they can be found in full-length in the following chapters. Those summaries serve the sole purpose to give the reader an understanding of how the work of the author fits in a broader context.

1.2.6 Research Aims

As outlined in the sections before, an in-depth understanding of the water oxidation reaction catalyzed by transition metal catalysts is of utter importance in order to further improve or develop novel WOCs. Here, computational chemistry bridges experiment and theory by modeling not only elusive catalytic intermediates and their properties but also by predicting how certain modification to the ligands might affect the catalytic activity. In this work we studied a variety of different WOCs for which we addressed a number of questions raised either by our experimental co-workers or ourselves. Besides that we also answered several methodical questions that arouse while simulating various properties of those WOCs. In this sense we always strove to find the best model system for the question at hand offering a reasonable compromise between accuracy and computational cost. Some specific research questions that were addressed in this work are highlighted in Sections 1.3.3, 1.3.4, 1.4.2 and 1.4.4).

1.3 Bioinspired Cobalt Based Water Oxidation Catalysts

1.3.1 The Cubane Structure in Co-based WOCs

In the following section the joint efforts made by Patzke, Luber and co-workers in developing and understanding Co^{II}-cubanes that are active WOCs will be discussed. It is not a coincidence that the groups have chosen to work on those structurally rather complex WOCs. Nature uses a cuboidal transition metal cluster known as oxygen-evolving-complex (OEC) to perform water oxidation. The OEC is embedded into a protein complex, the photosystem II (PSII), located in the thylakoid membrane of chloroplasts. The cluster itself is composed of calcium and three manganese cations bridged by oxygen atoms $\{CaMn_4O_5\}$. The fourth manganese cation, which is referred to as the dangling manganese, is located outside of the cubane core.[99]

The exact mechanism of the water oxidation reaction catalyzed by the OEC is still under debate, despite the numerous studies conducted over the past years.[100–108] Even though many attempts have been made to synthesize $CaMn_xO_x$ clusters that mimic the structure of the OEC, so far none of them has been able to oxidize water catalytically.[109–113] Nonetheless, the cubodial core structure has also served as an inspiration for the design of artificial WOCs that are not based on manganese such as for example cobalt-based cubanes.[38,114,115]

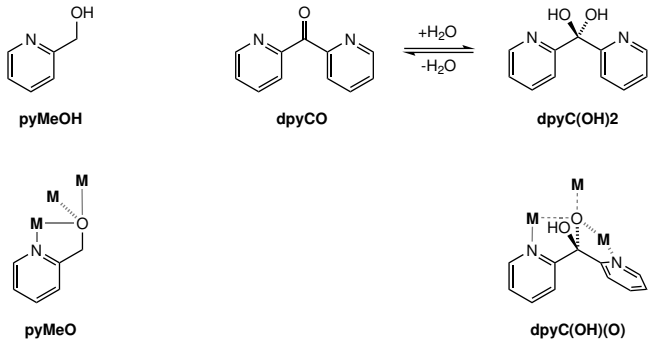


Figure 1.3: Chemical structure of the pyMeO = pyridin-2-ylmethanolato) and dpyC(OH)(O) = hydroxyldi-2-pyridinylmethanolato ligand, in their native form (top) and coordinated to several metal centers as in the Co^{II}-cubanes. Note how the hydroxide becomes part of the cubane core by coordinating three different metal centers.[38,115,116]

Among those bio-inspired WOCs are the previously mentioned Co^{II}-cubanes, that are not only unique in taking advantage of the lower oxidation state of cobalt (II instead of the usual III), which goes along with a more flexible ligand sphere i.e. faster ligand-exchange reaction rates, but also the use of ligands such as pyMeO = pyridin-2-ylmethanloate or dpyC(OH)(O) = hydroxyldi-2-pyridinylmethanolato (see Figure 1.3). The ligands do not only coordinate to the metal centers but are an integral part of the Co_4O_4 core.[38,115] The covalent linkage of the bridging oxygen to the ligands renders dissociation reactions highly unlikely. The linkage further prevents the bridging oxygen atoms from being protonated to participate in the oxygen-oxygen bond formation, which has been considered for model systems of Co_xO_3, where a $\{Co^{III}{}_4O_4\}$ cubane core is coordinated by aqua and hydroxido ligands.[117–120] We review the mechanistic studies of those model systems in detail in Chapter 3 and will therefore not discuss them in detail here.[13]

The main problem with Co^{III}-cubanes is their unknown structure in solution. Depending on the level of theory, and especially the solvation model applied, substantial differences in the stability (in terms of their total energy) of the large number of protomers can be observed. This in turn renders different reaction mechanisms and / or different intermediates energetically feasible or not. Beside the model systems for Cobalt-oxide WOCs there exists only one homogeneous WOC featuring a Co^{III} based cuboidal core, $\{Co^{III}{}_4O_4\}$ ($[Co^{III}{}_4(\mu^3\text{-}O)_4(\mu^2\text{-}OAc)_4(Py)_4]$, OAc = acetate, py = pyridine) which is often referred to as Dismukes-cubane.[114,121] Since those cubanes do not possess an aqua ligand, in a first step an acetate ligand has to change its coordination mode from μ^2 to μ^1 in order to allow the coordination of a water molecule.[119,122–124] This substitution reaction is expected to take place after a first

oxidation of the catalyst, and is followed by the deprotonation of the newly formed aqua ligand. This species might be part of the actual water oxidation cycle, however some authors argue that the experimentally supplied oxidation potential was not large enough to reach the later intermediates of the catalytic cycle.[122,124] Therefore, they proposed the association of a second hydroxide or water ligand to the cubane core with or without displacement of an acetate ligand. The existence of a dihydroxo species $\{Co^{III}{}_4O_4(OH)_2\}^-$ has been proven experimentally, yet considerations from thermodynamics led to the realization that the non-substituted $\{Co^{III}{}_4O_4\}$ species is indeed the predominant form in solution.[123]

It has also been shown that $\{Co^{III}{}_4O_4\}$ can be chemically oxidized to $\{Co^{III}{}_3Co^{IV}O_4\}^+$, this species has been isolated and was fully characterized. Due to the elusive nature of catalytic intermediates, their isolation and in-depth study tends to be very challenging. The oxidized species $\{Co^{III}{}_3Co^{IV}O_4\}^+$ was the basis for a wide range of kinetic and mechanistic studies. The most astonishing observation was that $\{Co^{III}{}_3Co^{IV}O_4\}^+$ evolves O_2 without the need of an additional oxidant, upon dissolution in basic aqueous media. From this Nguyen *et al.* concluded that $\{Co^{III}{}_3Co^{IV}O_4\}^+$ might disproportionate into $\{Co^{III}{}_3Co^{V}O_4\}^{2+}$ or $\{Co^{III}{}_2Co^{IV}{}_2O_4\}^{2+}$, since the existence of a formal $Co^V{=}O$ has been thought of to be a prerequisite for water oxidation.[123] The ability of polynuclear transition metal complexes to disproportionate in order to reach high oxidation states was also recognized in a novel proposal for the water oxidation mechanism by the nature's OEC.[125] In their latest study Nguyen *et al.* modified both the pyridine and the acetate ligands in order to alter the electronic properties of the cubane. The collected experimental data allowed them to derive an empirical (and specific to this kind of cubane) formula to predict reduction potentials of possible catalytic intermediates.[124] Using this tool, they revised their reaction mechanism proposed earlier. Their current mechanistic proposal suggests an active species with two hydroxido ligands in a *syn* configuration, as opposed to their previous active species which contained only one hydroxido ligand. The *syn* alignment of the hydroxido ligands renders an intramolecular reaction more feasible (i-WNA(OH)) than the earlier suggested intermolecular pathway where the O−O bond formation proceeds through a $WNA(OH_2)/WNA(OH)$ (see Figure 1.2). This is supported by recent *in situ* X-ray absorption spectroscopy measurements performed by Brodsky *et al.*, that provided the evidence for the existence of $\{Co^{III}{}_2Co^{IV}{}_2O_4\}^{2+}$.[126] This structure in principle also allows for an O−O bond formation by an intramolecular RC (I2M) mechanism.

Further studies on $\{Co^{III}{}_4O_4\}$ revealed that in the presence of permetalate, such as permanganate or perruthenate, a cobalt center is replaced by said metal ion forming a heterometallic cubane core.[127,128]. Spectroscopic data in combination with DFT simulations the $\{Co^{III}{}_3Mn^{IV}O_4\}$ and $\{Co^{III}{}_3Ru^{V}O_4\}$ cubanes suggest that the metal centers are strongly electronically coupled. Thus, they help to stabilize high oxidation states such as Ru^V. Further, for the first time a thorough characterization of a $Ru^V{=}O$ species was possible. This led to an experimental verification of the proposed radical character of the oxo ligand, together with the mechanistic implications discussed in Section 1.2.[127,128]

To summarize the work on Co^{III}-cubanes, the Dismukes-cubanes were studied in detail from an experimental and computational point of view. Despite the complexity of the system, the combination of both experimental and theoretical work led to a steady improvement of the proposed catalytic cycles. The extensive investiga-

tions also revealed, that some of the early synthetic protocols suffered from Co^{II} impurities, which were responsible for some of the catalytic activity. The problem was overcome by the introduction of an additional purification step in the synthetic protocol.[129,130] The observation that impurities can also contribute significantly to the overall catalytic performance highlights the importance of in-depth studies of the catalysts speciation in solution, as well as its stability.

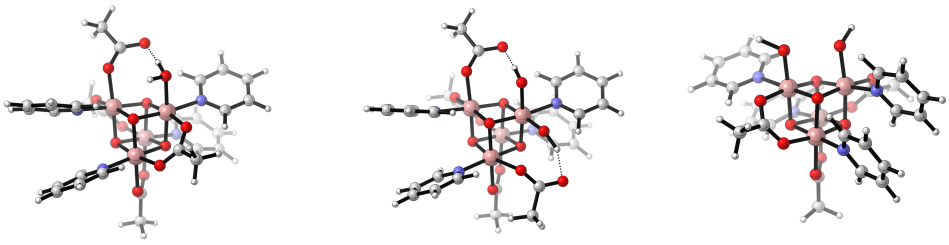


Figure 1.4: Schematic illustration of $[Co^{III}_4(\mu^3\text{-}O)_4(\mu^2\text{-}OAc)_4(Py)_4]$ where a single water molecule is associated (left[119]), or either two hydroxido ligands in a geminal- (middle[122]) or *cis*-configuration (right[124]). The following color code applies for all illustrations: *white* - hydrogen, *grey* - carbon, *dark blue* - nitrogen, *red* - oxygen, *green* - chlorine, *light pink* - cobalt, *lime* – erbium, thulium, ytterbium or holmium, *dark green* - ruthenium. Reproduced from Schilling and Luber.[13]

1.3.2 Discovery of Co^{II} Cubanes

The first Co^{II}-based cubodial WOC $[Co^{II}_4(pyMeO)_4(\mu\text{-}OAc)_2(\mu^2\text{-}OAc)_2(H_2O)_2]$ (pyMeO = pyridin-2-ylmethanolato) $\{Co_4(pyMeO)_4\}$ was discovered by Evangelisti *et al.*(see Figure 1.5).[115] Their turn-over-number (TON) was found to be similar to that of the Co^{III} cubanes while the turn-over-frequency (TOF) was determined to be two orders of magnitude higher. In addition, the catalytic performance was found to be pH dependent. Employing mildly basic conditions (pH = 8) resulted in the highest TON while even more basic conditions i.e. pH = 9 gave the highest TOF values. Those observations were in agreement with the spectroscopically determined pK_a value of the aqua-ligands of 8.7, indicating that at high pH values some of the aqua-ligands are already deprotonated. A secondary pK_a was found at 5.8 which was attributed to a potential substitution of an OAc ligand. Being aware of the issue of Co^{II} impurities that were found in the case of Co^{III}-cubanes, several experiments were conducted in order to exclude Co^{II}-leaching.[115]

Unfortunately, the isolation and characterization of elusive catalytic intermediates had turned out to be impossible in the case of the Co^{II} based cubanes. In particular since the catalytic activity was tested in a photocatalytic water oxidation setup, where the reaction solution contains many different species (e.g. buffer ions, persulfate ions, and photosensitizers). In order to elucidate possible mechanisms governing water oxidation catalyzed by $\{Co_4(pyMeO)_4\}$ Hodel and Luber carried out a theoretical study.[131]

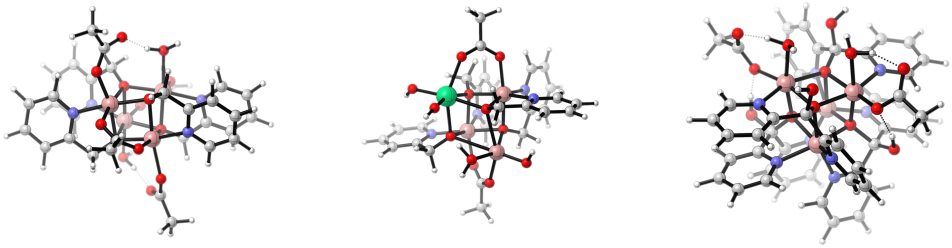


Figure 1.5: Schematic illustration of the cubanes studied by Patzke and co-workers: $\{Co_4(pyMeO)_4\}$ (left[115]) and $\{Co_3Ln(pyMeO)_4\}$ (middle[38]) (Ln = Er, Tm, Yb, Ho. Note: the structure shown is the model system used by Hodel and Luber,[67] and not the crystal structure reported by Evangelist *et al.*[38]), and the third generation $\{Co_4(dpyC(OH)(O))_4\}$ (right[116]). Reproduced from Schilling and Luber.[13]

Water Nucleophilic Attack Mechanism

The structural features of $\{Co_4(pyMeO)_4\}$, where the two cobalt-centers coordinated by aqua ligands are located on opposing faces of the cubane core, strongly suggest a WNA mechanism. If it is assumed that no proton- or redox isomerism can occur within the cubane core, then such a WNA mechanism would in essence be a single-site mechanism where the other three cobalt-centers merely act as spectators.

The mechanistic proposal by Hodel and Luber consists of four consecutive PCETs, during which the O–O bond formation takes place by a WNA onto the Co^{IV}=O species (see Figure 1.1). The largest difference in terms of free energy between two states was the transition from the S2 to S3 state. This chemical reaction includes the O–O bond formation by a WNA as well as a consecutive PCET forming a hydroperoxo species (S3) and is expected to be rate-limiting.[131] In order to release molecular oxygen, this species has then to be oxidatively activated i.e. it has to undergo a further ET and PT reaction. After which the association of a solvent molecule regenerates the catalytic ground state.

Since the thermodynamics can only give an incomplete picture of the catalytic cycle, Hodel and Luber completed the picture by modeling the relevant activation barriers. The barrier of the WNA was obtained by CI-NEB calculations.[43,44] This protocol depends on the premise that the reaction is a sole rearrangement of chemical bonds, i.e. the number of electrons and protons in the system have to remain the same. In the case of the O–O bond formation, this means that besides the Co^{III}–OOH species a H_3O^+ has to be formed. Even though Hodel and Luber used a generous first solvation shell, stabilization of a cation in close proximity to the catalyst appeared to be energetically unfavorable.[131] Note that this intermediate was not considered when evaluating the thermodynamics, since there the WNA is combined with a PCET. For this reasons the obtained activation barrier is likely an overestimate of the 'true' barrier.

Ligand Exchange Reactions

As mentioned earlier, there is also the possibility that an acetate ligand is substituted by either a water or hydroxide molecule. Alternatively, the ligand might also change its coordination mode from bidentate to monodentate upon association of a solvent

molecule. In order to assess the feasibility of ligand exchange reactions Hodel and Luber compared the energetics obtained by three different solvation models. Those models were: implicit (COSMO or COSMO-RS), static explicit solvation, as well as dynamic explicit solvation.[68] Here a dynamic treatment of the solvent and solute by the means of AIMD simulations is referred to as dynamic explicit solvation. Overall, the reactions energies obtained by AIMD simulations were qualitatively best reproduced by implicit solvation (COSMO-RS) calculations. This highlights the importance of hydrogen bonding that is in part accounted for in the COSMO-RS formalism but not in the standard COSMO implementation.[68]

Radical Coupling Mechanism

The extensive study on ligand substitutions in the first solvation shell suggested that $\{Co_4(pyMeO)_4(H_2O)\}^+$ ($[Co^{II}{}_4(pyMeO)_4(\mu\text{-}OAc)(\mu^2\text{-}OAc)_2(H_2O)_3]^+$), where a monodentate acetate is replaced by a solvent molecule, could exist under reaction conditions.[115] For $\{Co_4(pyMeO)_4(H_2O)\}^+$ there are several alternative routes for the O–O bond formation, among them is an intramolecular RC after four consecutive PCETs (see Figure 1.1), and an intermolecular WNA at the same stage.[131] Based on the lower activation barrier for the O–O bond formation as well as the more favorable thermodynamics Hodel and Luber concluded that the RC should be the dominant mechanism if the ligand substitution takes place.[131]

1.3.3 Co^{II}-Lanthanide Cubanes

The second generation of Co^{II} based cubanes features the same ligands as the first generation i.e. pyMeO and OAc. However, they include a redox-inert lanthanide ion $\{Co_3Ln(pyMeO)_4\}$ ($[Co^{II}{}_3Ln(pyMeO)_4(\mu^2\text{-}OAc)_4(\mu\text{-}OAc)H_2O]$, Ln = Er, Tm, Yb, Ho) (see Figure 1.5).[38] The catalytic performance of the $\{Co_3Ln(pyMeO)_4\}$ was found to be significantly improved compared to the one of $\{Co_4(pyMeO)_4\}$. This is surprising, since the redox-inert lanthanide ion is not expected to directly participate in the water oxidation mechanism. Like in the case of $\{Co_4(pyMeO)_4\}$ the catalytic performance of $\{Co_3Ln(pyMeO)_4\}$ was found to be strongly dependent on the pH and the buffer media. The best performance was usually measured under mildly basic conditions (pH = 8). It further has been shown that the aqua ligand undergoes rapid exchange with solvent molecules.[38] To some extent, the same is also true for the acetate ligands. This left Hodel and Luber with the question - what is the structure of the active species of the catalyst?

In a first step, Hodel and Luber determined the structure of the catalyst in solution, where they limited themselves to two out of the four lanthanide cations: Erbium which resides at the top in terms of catalytic performance and Thulium, which covers the lower end. AIMD simulations revealed that, in contrast to $\{Co_4(pyMeO)_4\}$, for $\{Co_3Ln(pyMeO)_4\}$ (Ln = Er, Tm) the displacement of bidentate acetate molecules by solvent molecules is thermodynamically favorable. Further, the displacement of a monodentate acetate coordinating the active cobalt center by a water ligand ($Co^{II}-H_2O$) as well as the substitution of both acetate ligands, coordinating in a bidentate manner to the lanthanide, by a hydroxide molecules turned out to be energetically feasible. Analogous observations were also made in static calculations employing implicit solvation. This led to the assumption that the dominant species in solution is $[Co^{II}{}_3Ln(pyMeO)_4(\mu^2\text{-}OAc)_2(OH)_3H_2O]$, rather than the pristine crystal structure.[38,67]

As in the case of $\{Co_4(pyMeO)_4\}$, the stability of the WOC in solution, before and after catalysis, was studied throughly by a wide range of experimental techniques.[38] All of which provided supporting evidence for the catalysts robustness.[67]

Possible Water Oxidation Pathways

Having established the catalytic resting state, Hodel and Luber further explored the thermodynamics and kinetics of the water oxidation mechanism. They used the same protocols like the ones that were already successfully applied in the case of $\{Co_4(pyMeO)_4\}$.[115,131] In the catalytic ground state only one cobalt center bears an aqua ligand which implies that a single-site mechanism has to govern the O–O bond formation. We refer to the cobalt ion that is coordinated by the aqua ligand as the 'active' cobalt center. This metal center is further coordinated by a hydroxido ligand in a *geminal* alignment with respect to the aqua ligand. This in principle allows for two reaction pathways, where the oxo ligand (Co^{IV}=O) is either on a different face of the cube as the lanthanide ion (pathway A) or on the same one (pathway B) (see Figure 1.6).

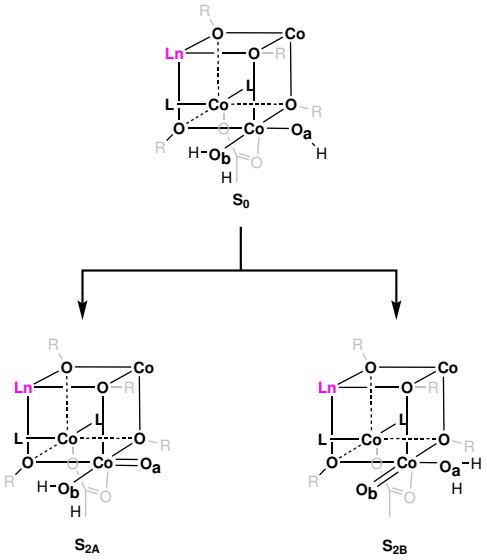


Figure 1.6: Schematic illustration of selected states of the two possible pathways A (left) and B (right); For the sake of clarity all ligands on the other cobalt or lanthanide cations are not shown, except for the one where the water oxidation takes place.

Hodel and Luber compared the thermodynamics of the two pathways obtained using explicit static solvation. In general, the thermodynamics of both catalysts was reported to favor pathway A over pathway B. However, in the case of the S2 state the energetics of the two catalysts differentiated significantly. While $\{Co_3Er(pyMeO)_4\}$ clearly favored pathway B over pathway A, the opposite was true in the case of $\{Co_3Tm(pyMeO)_4\}$.[67] This was attributed to a distortion of the cubane core, caused by the 'active' cobalt atom which is slightly pulled out of the cubane cage upon reaching the S2 state. This opening was found to be thermodynamically favorable. They also calculated the activation barriers for the O–O bond formation by WNA using CI-NEBs. For both catalysts pathway B turned out to be energetically favor-

able.[67] A likely explanation for this observation is the protonation of the hydroxido ligand instead of the formation of a hydronium as found in the case of pathway A.

Opening of the Co^{II}-Lanthanide Cubane Cage

Research Aim - How important is the 'open' conformation of Co^{II}-based cubanes for the catalytic cycle?

The potential co-existence of two distinctive conformers raised the question whether those 'open' conformers were thermodynamically and kinetically accessible, and whether their participation in the catalytic cycle would be beneficial (1). Besides that, we also started to question the reliability of the explicit static solvation when it comes to assess the relative stability of conformers (2).

Inspired by those distorted cubanes, we identified structures where the 'active' cobalt center is even further pulled out of the cubane core (for the full-length story of the 'open' cubane see Chapter 4). We will refer to those structures as 'open' (see Figure 1.7). In fact, the 'active' cobalt center obtains a tetrahedral coordination as opposed to an octahedral one as in the case of the 'closed' cubane. Further, one of its ligands forms an intramolecular hydrogen bond with the alkoxide of the neighboring ligand, thereby stabilizing the 'open' form. This is of course only possible for pathway A, where the active ligand is not on the same face of the cubane as the lanthanide ion. We located those 'open' cubane structures for both lanthanide ions, employing either implicit solvation or explicit static solvation, for all proposed intermediates of the catalytic cycle (S0 to S4). Employing implicit solvation we determined for both catalysts the 'open' form of the S0 state to be slightly favored in terms of free energy, while the two forms are isoenergetic in the S1 state. For the S2 state, which is of great importance for the reactivity of the catalyst, we found $\{Co_3Er(pyMeO)_4\}$ to prefer an 'open' configuration while $\{Co_3Tm(pyMeO)_4\}$ favors 'closed' configuration. Intermediates of the subsequent states favor a 'closed' form. Those states are further stabilized by an additional intramolecular hydrogen bond between the hydroperoxide and the hydroxido ligand of the 'active' cobalt center. Looking at thermodynamics of the whole catalytic cycle we found no overall benefit of an all 'open' vs. an all 'closed' cycle. We also calculated the barrier of the interconversion (open-closed) to address the question whether the two species might co-exist. The approximated barrier of this transition was obtained by self-consistent optimization of the reaction path.[42] The small barrier height ($< 6\,\mathrm{kcal\,mol^{-1}}$) rendered the 'opening' of the cubane core under experimental conditions possible. Based on those observation we suspect the 'open' form of the cubane to only play a role for the initial states of the catalytic cycle. This is to some extent reminiscent of one of the proposed catalytic mechanisms of the OEC, where the cubane core possesses both an 'open' and a 'closed' form in the S2 state. The 'open' form is then expected to undergo a further PCET to reach the S3 state.[132] The final steps of the cycle, including the O–O formation are still under debate.[125]

Further, employing explicit static solvation we found the 'open' structure to be energetically favored throughout the catalytic cycle. Comparing the 'open' and 'closed' cubane structures with and without explicit solvation we were able to verify that intermediates S0 and S1 indeed favor an 'open' configuration while the stabilization

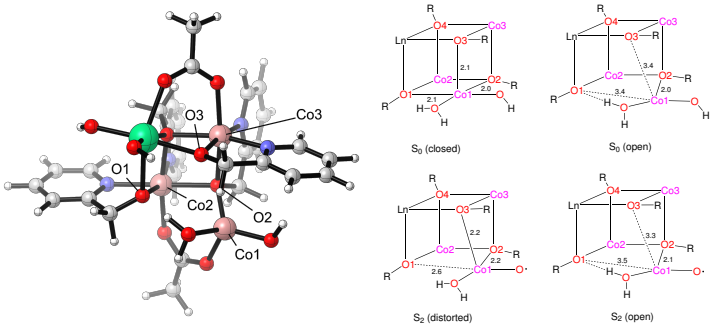


Figure 1.7: Ball-stick representation of the 'open' form of $\{Co_3Ln(pyMeO)_4\}$ (left). Schematic drawing of the 'closed' and 'open' form discussed in this section, as well as the 'distorted' form already mentioned in the context of the mechanism in the previous section.[67] Reproduced from Schilling *et al.*[133]

of the later intermediates was attributed to changes in the solvation shell. The fact that there are many different energetic minima with respect to the solvation shell makes the calculation of energy differences between catalytic intermediates somewhat arbitrary. The reason for this is the fact that it is uncertain whether the energy difference is caused by the change of the solute or primarily by a structural change of the solvent shell, which might occur independent of the solute. The introduction of many explicit solvent molecules into the model system thus in principle requires the use of sampling techniques such as MD.

1.3.4 A New Generation of Co^{II}-Cubanes

We tackled this problem in our latest study on Co^{II} based cubanes i.e. $[Co^{II}_4((dpyC(OH)(O))_4)(OAc)_2(H_2O)_2]^{2+}$ (dpyC(OH)(O) = hydroxy-di(pyridin-2-yl)methanolato, $\{Co_4(dpyC(OH)(O))_4\}$) (see Figure 1.3).[116] Those clusters are the basis of the third generation of cubanes developed by Patzke and co-workers. A major difference compared to the previous generations is the use of a tridentate ligand, instead of the bidentate ones used before. The novel ligand might be thought of as an extended version of pyMeO, where the methylene hydrogen atoms were replaced by a pyridine and a hydroxide. The similarity to the previous ligand pyMeO was intended in order to take advantage of the favorable properties observed for $\{Co_4(pyMeO)_4\}$ and $\{Co_3Ln(pyMeO)_4\}$, such as the incorporation of the ligand into the cubane core, which prevents the bridging oxygens from participating in the water oxidation mechanism. The $\{Co_4(dpyC(OH)(O))_4\}$ cubanes are not only interesting from a mechanistic point of view, but they also possess another key feature unknown for the other cubanes; it is experimentally possible to stoichiometrically replace Co^{II} by Ni^{II}, while keeping the same structure. This potentially allows for unprecedented insight into the catalytic mechanism, in particular on how the nature of the transition metal might alter its kinetics and/or thermodynamics.

There is another important structural feature: For the first time two aqua ligands are located on the same face of the cubane. Together with the non-coordinating hydroxyl group of the ligands, they form an edge-site which strongly resembles the solvated surface of Co_xO_3. This makes $\{Co_4(dpyC(OH)(O))_4\}$ the ideal model system to study both homogeneous and heterogeneous reaction mechanisms for

water oxidation, although the catalytic performance was found to be slightly worse than the previous generations (TON: 20, TOF 0.24 s^{-1}).[116] Those two aqua ligands have a crucial impact on the possible reaction mechanism. Since these ligands are coordinated to metal centers with the same chemical environment, it is likely that both of them are involved in the water oxidation mechanism. Therefore it remains unknown whether the reaction follows a WNA(i-B assisted), i-WNA(OH) or an i-RC pathway to form the O–O bond (see Figure 1.2). In contrast to this, for the previous generations of cubanes generally a single-site mechanism was assumed unless a second water molecule is expected to associate to the cubane core.[67,131]

Again, it turned out to be difficult to characterize catalytic intermediates experimentally. However, recording X-ray absorption near edge structure (XANES) under catalytic conditions revealed the partial oxidation of the cubane and its slow recovery to the catalytic ground state, i.e. the complete reduction to an all Co^{II} state. Even though those experiments cannot give an insight into the reaction mechanism, they strongly support the claim that the catalyst is stable under catalytic conditions.

Structural Flexibility of the Catalytic Ground State

Research Aim - What is the dominant cubane species in solution?

Studying earlier generations of Co^{II}-based cubanes made us aware of the combinatorial challenge that could arise from the speciation of $\{Co_4(dpyC(OH)(O))_4\}$ in solution. This raised two important questions: (1) does the cubane core also 'open' under ambient conditions and (2) is the dpk ligand involved in proton reshuffling once the catalyst is deprotonated.

As a prerequisite for a mechanistic proposal, the species in solution had to be determined. We did so by performing AIMD simulations of the cubane that has been solvated by either 208 or 452 water molecules (see 2.5.2 for computational details). Being curious about the discovery of the 'open' structures for $\{Co_3Ln(pyMeO)_4\}$, we monitored the Co–Co distances as well as the volume enclosed by the cubane cage during the AIMD runs. While we observed significant fluctuations of all Co–Co distances, some of which are correlated with each other, we did not detect an opening of the cubane cage, probably due to geometrical restraints imposed by the bigger ligands which hold the cobalt ions in place. Since the highest catalytic activity was observed under mildly basic conditions (pH = 8.5), which is also the sole experimentally determined p$K_a$value of the WOC, we also checked whether the other functional groups of the 'edge-site' could engage in proton transfer reactions. Modeling of a deprotonated cubane ($[Co^{II}_4((dpyC(OH)(O))_4)(OAc)_2(H_2O)(OH)]^+$) revealed the involvement of not only the ligands themselves but also of the solvent molecules in various proton transfer reactions, resulting in deprotonation of the hydroxyl group of a ligand in favor of the aqua ligand. This study highlights the importance of an appropriate solvation model. At the same time it also complicates the proposal of a catalytic mechanism, since there are many protomers for each intermediate. Further investigations are currently ongoing in the Luber group.

Controlling the Self-Assembly Mechanism of Cubanes

In the meantime Patzke and co-workers further investigated the self-assembly mechanism of Co^{II}-based cubanes incorporating the dpyC(OH)(O) ligand.[134] In an exhaustive study they varied the parameters of the synthetic protocol i.e. the metal source, the nature of the counter-ions, as well as their respective concentrations. By doing so, they were able to synthesize a library of $\{Co_4(dpyC(OH)(O))_4\}$ analogues. Based on the coordination mode of the dpyC(OH)(O) ligand two distinctive types of cubanes were identified (see Figure 1.8). This suggested either a cobalt aqua ion or a cobalt acetate complex as the initial building-block of the cubane. Those experimental findings, the thermodynamics of various cobalt-complex formation reactions were modeled using a static-DFT approach (see Section 2.5 for computational details). Together with the experimental findings those simulations supported the hypothesis of different assembly pathways for the two types of cubanes. However, there are still various open questions such as the influence of the ionic-strength on processes such as nucleation and ligand association/dissociation reactions.

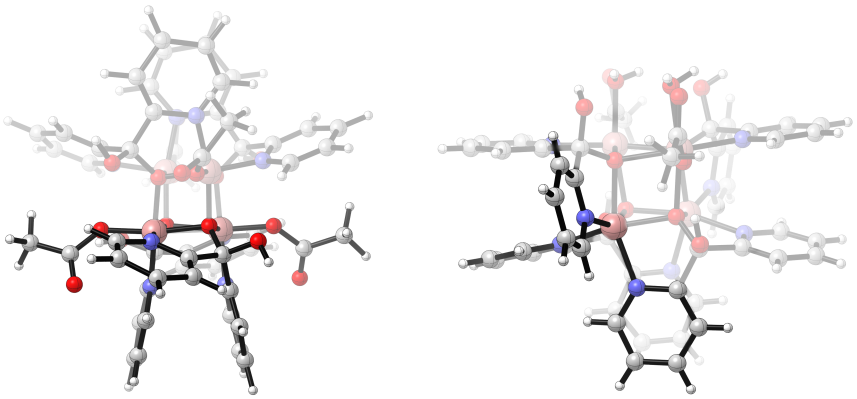


Figure 1.8: Ball-stick representation of the $[Co^{II}_4((dpyC(OH)(O))_4)(OAc)_4]$ cubane, known as 'type 1' (left) and of the previously studied 'edge-site' cubane ($[Co^{II}_4((dpyC(OH)(O))_4)(OAc)_2(H_2O)_2]^{2+}$) now termed as 'type 2' (right). The key features of both structures are highlighted: In case of 'type 1' two identical dimers form the cubane core, while for 'type 2' cobalt centers coordinated by three pyridyl groups, each part of a different dpyC(OH)(O) ligand, form a bowl-shaped cavity around the cubane core. Reproduced from data published by Song *et al.*[134] (*unpublished figure*).

Besides Co-based WOCs, we have also studied mononuclear Ru-WOCs in collaboration with experimental groups. Those WOCs possess pyridyl based ligands, which are described in more detail in the following. As in the case of the Co^{II}-cubanes, the contributions of the Luber group to the field are presented in a broader context. In this sense collaborative work with contributions of the author are discussed in detail. Additionally, the works where the author is the main contributer are briefly summarized and some of the important findings are highlighted. The full-length studies for those works can be found in Chapter 5, 6 and 7.

1.4 Mononuclear Ru-based Water Oxidation Catalysts

Noble metals such as ruthenium have been widely used in catalysis for water oxidation, notably since the first discovered molecular WOC, the so called 'blue-dimer', was a ruthenium based transition metal complex.[135,136] Since then, there have been numerous publications on different catalysts and their possible mechanisms. A full review of the field is beyond the scope of this work, therefore we want to point the reader's attention towards some important reviews.[22,137–143] In the following sections we will primarily focus on computational work done in the Luber group. However, we put our work into a broader context, in order to embed our contribution into the work of others in this field.

1.4.1 Ru-based WOCs bearing the bda Ligand

The most common Ru-based WOCs feature a bda (2,2′-bipyridine-6,6′-dicarboxylic acid) ligand and have been studied by various groups. For instance Tong *et al.* were able to show in a combined experimental and theoretical study how introducing rigidity into the ligand, i.e. the usage of pda (phenantroline-2,9-dicarboxylic acid) instead of bda, led to a different reaction mechanism for the O−O bond formation (see Figure 1.9).[34] While the flexible bda ligand favored an oxo-oxo coupling pathway, the rigid pda ligand preferred a WNA mechanism. However, altering the reaction mechanism had a profound effect on the catalytic performance, which was in the case of the pda WOC significantly lower than the bda containing catalyst in terms of TONs i.e. 310 vs 1200 for [Ru(L)(pic)$_2$] (where L is pda rsp. bda, pic = picoline). Despite the lower TONs an enhanced lifetime of pda bearing catalysts was observed. However, it remains unclear whether the increased lifetime is directly translated into an enhanced stability or if the water oxidation reaction itself is just slowed down. In an attempt to rationalize the effect of the rigidity, the reorganization energy of the ligand was calculated i.e. the energy difference required to deform the free (non coordinated) ligand towards its geometry, when coordinated to the metal center.[34] Independent of the oxidation state of the metal center, the reorganization energy remained the same for the pda ligand, while for the bda ligand drastic changes (up to 10 kcal mol^{-1}) were observed. This highlights the fact that the bda ligand can adopt to the smaller ionic radii of the higher oxidation states while there is a growing mismatch between the pda ligand and the metal ion upon oxidation.

Taking all those observations into account, our experimental co-workers in the groups of Alberto and Llobet came up with a ligand which in essence is the exact opposite of the pda ligand. Instead of locking the bipyridine subunit into one plane, they enforced non-planarity by replacing pyridine by isoquinoline resulting in biqa ([1,1′-biisoquinoline]-3,3′-dicarboxylic acid) (see Figure 1.9).[35] While this ligand can in principle rotate along its 1,1′-bond and therefore adopt to changing ionic radii easily, there is a significant penalty associated in case the dihedral angle describing the alignment of the two isoquinoline subunits approaches small values, i.e. planarity. To complete the catalyst, they incorporated pic and 6-bromoisoquinoline (isoq) as axial ligands. The rational behind this is an outstanding study on Ru-bda WOCs by Duan *et al.* where they reported a drastic increase in catalytic performance upon replacing pic with isoq, in particular the TOF increased by almost an order of magnitude (303 s^{-1} compared to 32 s^{-1}).[144] The reason for this increase is

suspected to be the favorable interaction of the axial ligands in the I2M mechanism, stabilizing either the encounter complex or the TS. Elucidating the influence of the axial ligands has since then received a lot of attention, and several groups tried to explain its nature from both an experimental and theoretical point of view.[74,145–149] While the main focus in the early 2010s was primarily the interaction between the axial ligands, nowadays the attention shifted towards the influence the solvation shell and more and more elaborate computational models are applied,[33,71,72,150,151] where the solvent is not only described by an implicit solvation model but rather fully atomistic. To do so, simulations based on, for instance, the EVB model were used.[71,72]

There also have been attempts to computationally design better WOCs by modifying their ligand framework. In the following we briefly highlight a few studies where specific properties such as the activation barrier of the O–O bond formation or the coordination strength of ligands have been tuned.[146,152–158] In 2012, Duan *et al.* conducted a DFT-screening approach of axial ligands to overcome the tendency of the axial ligands to dissociate.[146] Correlating the free energy of the ligand-solvent exchange reaction with the highest occupied molecular orbital (HOMO) energy of the ligand allowed them to identify some potential candidates for tightly bound axial ligands which then were synthesized and experimentally tested. Modifications to the bda ligand were later studied by Kang *et al.* where they compared the barriers for the WNA and the I2M mechanisms.[154] They not only replaced the carboxylates with amides but also pyridyls with phenyls and N-heterocyclic carbenes (NHCs). The synthetic applicability of some of those modifications might be questionable, however they observed some general trends in terms of barrier heights. In a unique study, the effect of electron withdrawing and donating substituents on the metal coordinated pyridine derivatives was investigated by Asaduzzaman *et al.*[155] They found that by tuning the electron density at the metal center of a $Ru-O_2$ species the most stable spin state could be varied. This is crucial since the catalysts are expected to evolve triplet oxygen. Stabilization of a triplet state for $Ru-O_2$ would therefore mean a facile O_2 liberation. Richmond *et al.* investigated the influence of modifications to the axial ligand on the activation barrier of the O–O bond formation according to the I2M mechanism.[158] They found the introduction of fluorinated substituents in the *para* position of the axial ligand to not only lower the activation barrier but also to be thermodynamically favorable. This is rationalized by a decrease of the non-covalent interactions of the axial ligands.

1.4.2 Modifying the bda Ligand

Research Aim - How to explain the different O_2 evolution rates?

For bda derived WOCs it has been shown that the exchange of the axial pic ligand by an isoq ligand led to drastic increase in O_2 production. This was attributed to a reduction of the activation barrier for the I2M mechanism.[144] As the biqa and bda ligands are closely related, it is reasonable to ask whether the same explanation also applies for the biqa derived WOCs.

The biqa derived WCOs, where either pic or isoq are used as axial ligands, were studied by our experimental collaborations in a chemical water oxidation setup

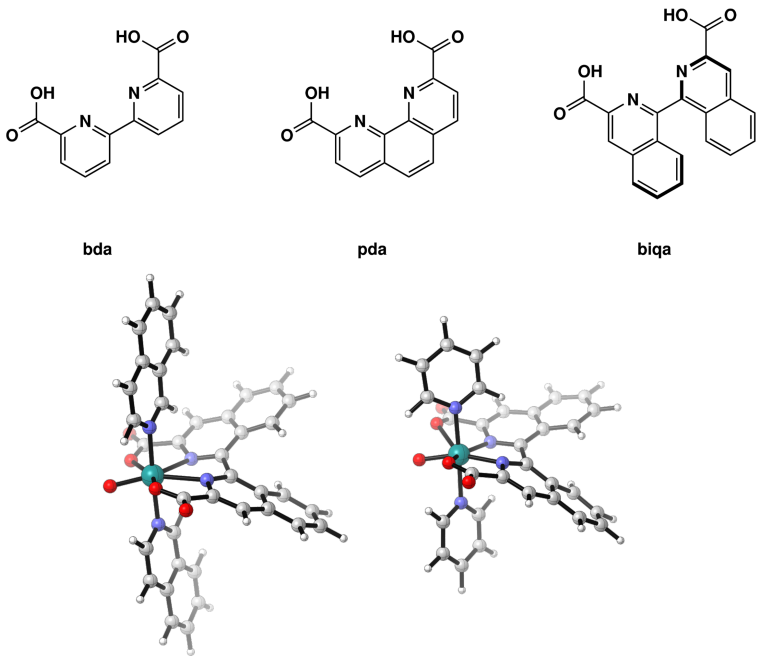


Figure 1.9: Schematic representation of pyridine based bicarboxylic acid ligands used for WOC of the type $[Ru(L)(R)_2]$, where L = bda, pda or biqa and R = pic or isoq (top). Note, in the simulations pic was replaced by py and 6-bromo-isoq by isoq. Ball-stick representation of $[Ru(biqa)(isoq)_2]$ and $[Ru(biqa)(pic)_2]$ catalysts investigated by Scherrer *et al.*(bottom).[35] Reproduced from geometrical data published by Scherrer *et al.*[35]

using CAN as a sacrificial oxidant. They found the WOCs to obey a first-order rate-law with respect to the catalyst concentration.[35] This implies that a WNA rather than a I2M mechanism governs the O−O bond formation. Surprisingly, the initial rates (0-4 s) of the two catalysts were determined to be vastly different, while the TOFs for the remaining period (4-20 s) were almost identical ($0.63\,s^{-1}$ (pic) and $0.57\,s^{-1}$ (isoq)). Those values were similar to the ones reported for Ru-bda-pic measured under the same conditions.[34] This is reminiscent of a study on Ru-bda-pic/isoq conducted by Duan *et al.* the initial TOF of the isoq bearing catalyst was significantly higher than the one with pic as axial ligand.[144]

In order to understand this behavior we conducted a DFT study employing standard protocols including implicit solvation by using COSMO (see Section 2.5 for computational details). As for the Ru-bda complexes,[159] we assumed a sevenfold-coordinated Ru^{V}=O species as the key intermediate. The barriers for the WNA mechanism were calculated for both the biqa and the bda ligand, which showed that the influence of axial ligands on the barrier of the WNA mechanism is minimal. As already observed by others, a carboxylate acted as an intramolecular base during the reaction mechanism, stabilizing the formed hydroperoxo-species.[34] Likewise we observed that the introduction of an additional water molecule acting as a proton shuttle between the nucleophilic water and the carboxylate lowers the barrier. This finding might be somewhat artificial, since - in the absence of solvent the stabilization of a hydronium in close proximity to the positively charged catalysts

is energetically unfavorable, especially if the ion is not stabilized by a sufficiently large hydrogen bond network. The barriers for the WNA were then compared with the ones for the I2M mechanism (see Figure 1.10). Those barriers were obtained by performing scans along the constrained O–O bond. Due to the bulkiness of the catalysts there were several conformers possible with various degrees of $\pi - \pi$ stacking. This was further complicated by the fact that the two $^{1}[Ru^{V}{=}O]$ species could obtain an (open shell) singlet or triplet spin state. As a consequence the barriers for the O–O bond formation have to be interpreted with caution. In the case of the pic ligand, the activation barrier of the I2M mechanism was found to be larger than the one of the WNA, which is in agreement with the first order rate constant. On the other hand, in the case of the isoq ligand we found that the barriers for the I2M mechanism and WNA are quite similar, and therefore do not allow for a unambiguous assignment of a reaction mechanism. This is in agreement with the intermediate rate constant observed in experiments. Analysis of the post-catalytic solutions by high performance liquid chromatography revealed for both catalysts free biqa ligand in solution. This suggests that the catalysts were unstable under reaction conditions and potentially deactivated after a few turn-overs. Those observations also raised the questions regarding what is the 'true' active species, and whether it is the same for both catalysts.

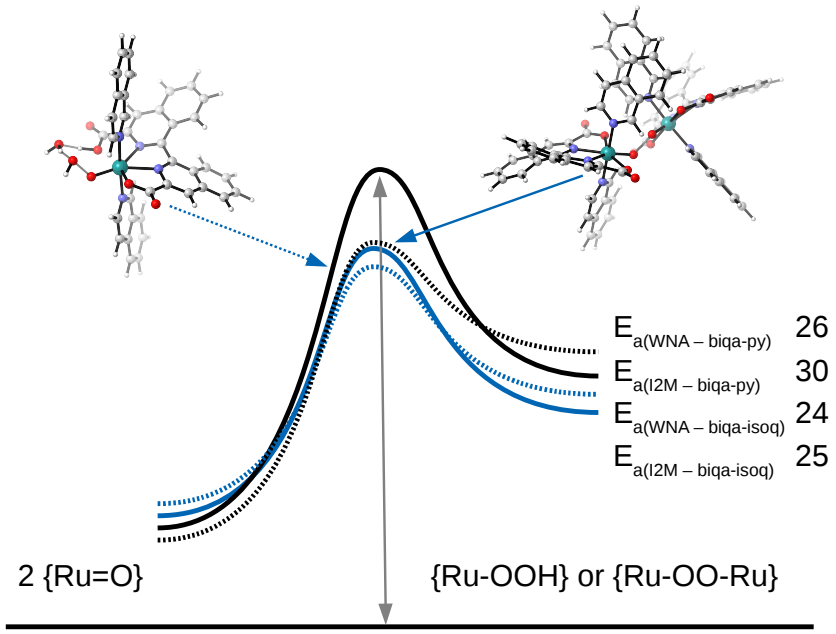
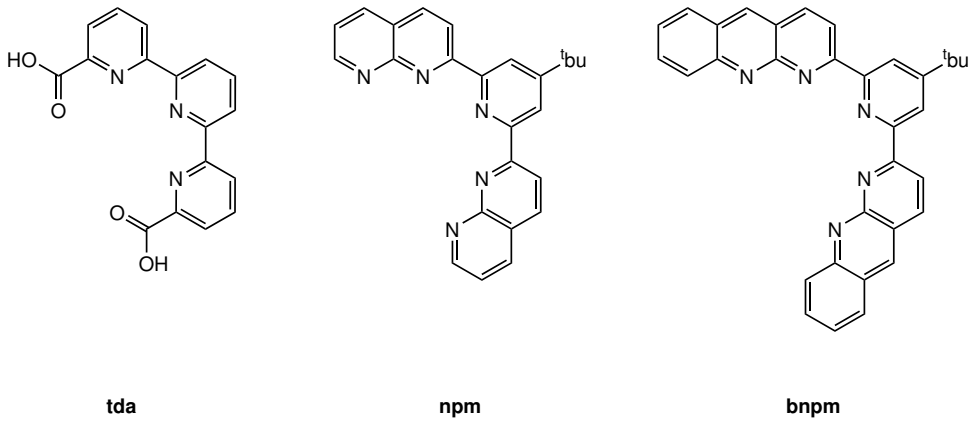


Figure 1.10: Electronic activation energies (kcal mol^{-1}) for the WNA and I2M mechanism of [Ru(biqa)(R)$_2$], where R = py (dashed lines) or isoq (solid line). Two exemplary TS structures are shown, the color and line style are the same as the energy profile. The activation energies are ordered by the stability of the respective product. Note in the case of the WNA mechanism, there is an additional water molecule which acts as a proton shuttle. Reproduced from data published by Scherrer *et al.*[35]

1.4.3 More Complex Polypyridine Ligands

Another type of ligand modification has received a lot of attention, namely the introduction of an additional pyridine subunit into the bda ligand,[36] resulting in the tda ligand ([2,2′:6′,2″-terpyridine]-6,6″-dicarboxylate) (see Figure 1.11). A complete recapitulation of all the studies related to Ru-tda catalysts is beyond the scope of this work and we refer the reader to corresponding publications by Matheu *et al.*[36,160,161] Nevertheless, we want to point out some of their characteristic features: First, the first-order kinetics in terms of the catalyst concentration which suggests a WNA mechanism for the O−O bond formation; secondly, flexibility of the tda ligand which can change its coordination mode and therefore adopt to changing ionic radii; thirdly, the decoordination of a carboxylate, which enables it to act as an intramolecular base and thereby potentially facilitates a WNA.[36] In two recent studies the role of this carboxylate was investigated by means of AIMD and standard DFT protocols,[37] where the emphasis focused on the pH dependent mechanism.[162]

Beside the carboxylate bearing ligands, there is also a huge number of Ru-based catalysts which possess polypyridine ligands or derivatives thereof. Among them is a family discovered by Thummel and coworkers, the mononuclear version of which contains npm (4-t-butyl-2,6-di(1′,8′-naphthyrid-2′-yl)-pyridine) as well as two pic in the axial position (see Figure 1.11).[163–169] The npm ligand contains five basic nitrogen atoms which potentially could coordinate to the metal center, however the geometrical constraint of the naphthyrid-subunit in fact limits the denticity of the npm ligand to three. As a consequence, the metal ion is not completely locked inside the ligand pocket and, as for the bda catalysts, decoordination of axial ligands might be possible deactivation routes. A possible approach to overcome this is the usage of a higher denticity, such as pentadentate tetrapyridine ligands introduced into the field of Ru-based WOCs by Zhao and cowokers.[170,171]However, those ligands include ethylene linkages which potentially are susceptible to oxidation under catalytic conditions.



tda **npm** **bnpm**

Figure 1.11: Chemical structures of some selected pyridine based ligands, commonly used in WOC catalysts: tda ([2,2′:6′,2″-terpyridine]-6,6″-dicarboxylate) (left),[36] npm (4-*t*-butyl-2,6-di(1′,8′-naphthyrid-2′-yl)-pyridine) (middle),[163] bnpm (2,6-bis(benzo[*b*]-1′,8′-naphthydridin-2′yl)4-*t*-butylpyridine) (right).[166]

1.4.4 Pentapyridine based Ru-WOCs

In 2017, Gil-Sepulcre *et al.* overcame this shortcoming with the development of the pentadentate ligand Py5R (6,6″-(R(pyridin-2-yl)methylene)-di-2,2′-bipyridine, where R is either a methoxy- or methyl group) (see Figure 1.13). The main framework consists of two bipyridines and a pyridine covalently linked at a tertiary carbon.[172,173] On the one hand this ligand has the great advantage that the axial pyridine is covalently linked and therefore cannot diffuse away upon decoordination. On the other hand, the ligand restricts the accessibility of the Ru-center and therefore potentially prevents the formation of a seven-fold intermediate which was noticed to be advantageous for many other Ru-based catalysts.[159] Surprisingly it was found that the catalytic performance of $[Ru(Py5OMe)(OH_2)]^{3+}$ obtained by a ligand exchange reaction from $[Ru(Py5OMe)(Cl)]^{2+}$ was inferior to its precatalyst $[Ru(Py5OMe)(Cl)]^{2+}$ (TOF: $0.037\,s^{-1}$ vs. $0.710\,s^{-1}$). Those observations, together with a previous study on the *cis* / *trans* isomerization of the ligand,[172] led to the conclusion that the pyridyl subunit is indeed able to decoordinate and has to be replaced by an aqua ligand in order to be able to catalyze water oxidation.

Comparing the Py5OMe and Py5Me Derived Catalysts

Research Aim - What makes the difference between Py5OMe and Py5Me derived catalysts?

In simple terms differences in catalytic activities are always related to either the electronics or the sterics of the transition metal complex. Those properties refer to the electron density at the metal center and the accessibility of the reactive site. However, how can differences in the catalytic activities be explained when the modifications of the ligand neither directly alter the electronics of the metal center nor are in proximity to the active site?

Unexpectedly, the catalytic performance of the two catalysts with the very similar ligands Py5OMe and Py5Me turned out to be rather different. At first glance this is difficult to rationalize since the different functional groups are neither directly coordinated to the metal center nor part of a conjugate system and therefore hardly alter the electronics of the metal center. There is also no significant steric difference, in particular with respect to the suspected WNA mechanism. In order to verify this hypothesis we conducted a full-scale DFT study where we compared the two catalysts in terms of their thermodynamics and kinetics (see Section 2.5.1 for computational details). Investigating the differences in free energy relating to possible intermediates of the catalytic cycle, the thermodynamics of the whole catalytic cycle were comparable among the two catalysts. Differences if there were any, were within the error margin of DFT, except for the substitution of the pyridine with a water molecule - either at the Ru^{II} or Ru^{III} stage.

The decoordination of the pyridine followed by an association of solvent molecule was found to be thermodynamically more favorable in the case of the Py5OMe WOC. We did not attempt to calculate a TS for this reaction, due to the unknown nature of the reaction (concerted vs. step-wise, a possible involvement of solvent molecules etc.). Instead, we focused our attention on the barrier of the WNA forming the O–O bond. This process is especially interesting for the WOCs at hand, since they possess

a decoordinated pyridine in close proximity that is restricted in its conformational freedom due to its covalent linkage to the metal coordinated bipyridyl fragments. The pyridine may thus likely act as proton acceptor that facilitates the WNA by deprotonating the approaching water molecule. This was confirmed by our calculations which revealed a concerted mechanism for the O–O bond formation and the concurrent deprotonation of the nucleophile by the pyridine (see Figure 1.12).

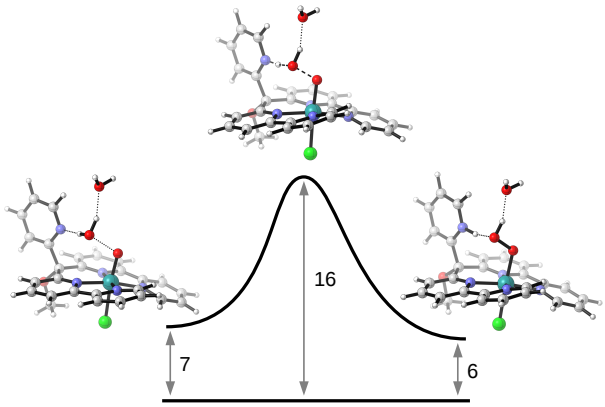


Figure 1.12: Reaction scheme of the O–O bond formation for a base assisted WNA of $[Ru(Py5OMe)(Cl)]^{2+}$. Reproduced from Gil-Sepulcre *et al.*[173]

The barriers for the two catalysts were, like their thermodynamics, very similar, i.e. 16 kcal mol^{-1} (Py5OMe) vs. 14 kcal mol^{-1} (Py5Me) in terms of absolute barriers - relative to the Ru^{V}=O species and two water molecules at infinite distance. The differences were even smaller when comparing barriers relative to their respective associated reactants (ARs), i.e. an encounter complex of Ru^{V}=O species and two water molecules. In this case the barrier for both catalysts was found to be 9 kcal mol^{-1}. Those barriers suggest that the O–O bond formation is not the distinguishing factor for the different catalytic activities of the catalysts featuring Py5OMe and Py5Me.

In an exhaustive kinetic study our experimental co-workers determined several rates of ligand exchange reactions such as the chlorido-water and the pyridine-water exchange (see Figure 1.13). Most ligand exchange rates were similar for both ligands. An exception was the chlorido-water exchange rate, which was found to be an order of magnitude larger in the case of the Py5Me catalyst. The faster chlorido exchange therefore suggests that the Py5Me deactivates more quickly than the Py5OMe species which results in an overall poorer catalytic performance. In a follow-up study we further investigated the role of chlorido-water exchange by calculating TSs for a concerted reaction mechanism.[174] For more details see Chapter 5. We found that - independent of the oxidation state of the metal center - the (absolute) barriers in the case of the Py5Me catalyst were about 4 kcal mol^{-1} smaller than the corresponding ones for Py5OMe bearing WOC. This observation is in qualitative agreement with the experimental measurement.[173] Those differences might be caused by the sterical clash of the substituent at the tertiary carbon with the two bipyridine subunits. Those interactions are smaller in the case of OMe since the oxygen atom acts as a spacer between the methyl group and the rest of the ligand.

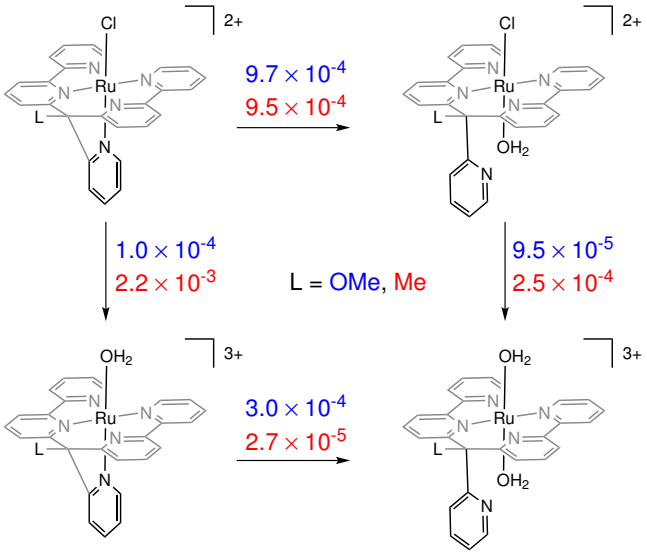


Figure 1.13: Rate constants for ligand dissociation / association reaction, either experimentally measured or obtained from kinetic modeling of the Py5OMe (blue) and Py5Me (red) bearing catalysts. All rate constants are given in s^{-1}. Adopted from Gil-Sepulcre *et al.*[173]

Another potential bottleneck in the catalytic pathway might be the dissociation of the pyridine ligand. Since the nature of this reaction remains unknown (i.e. concerted, step-wise, solvent assisted etc.) we decided to conduct dissociation scans of the $N_{Py}-Ru$ bond as well as of the H_2O-Ru bond. The decoordination of the pyridine appears to be favorable in the case of the Py5OMe ligand while the dissociation of the water ligand is less favorable (see Figure 1.14). Those scans are in agreement with calculations of the thermodynamics which rendered the displacement of pyridine by water more favorable in the case of Py5OMe.

To summarize the results, the difference in catalytic performance between the two catalysts featuring Py5OMe and Py5Me cannot be rationalized by simply looking at the water oxidation mechanism. As suggested by both experimental results as well as DFT calculations, side reactions, namely the chlorido-water exchange followed by a pyridine-water substitution, lead to a trans-diaqua complex which presumably is inactive for water oxidation. Further the energetics of the formation of the active catalyst, i.e. the pyridine-water exchange, is also different for the two ligands. A combination of those two factors, the faster rate of a reaction leading to the deactivation of the catalyst as well as a potentially slower rate to form the active catalyst, leads to the overall inferior performance of the catalyst featuring Py5Me compared to the one with Py5OMe.

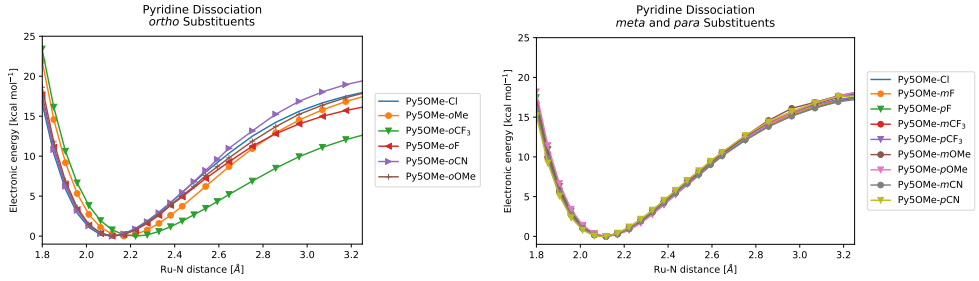


Figure 1.14: Energy profiles of pyridine dissociation reactions for different substitution patterns at the pyridine of [Ru(Py5OMe)]. Reproduced from Schilling *et al.*[174]

Rational Ligand Design

Research Aim - Can we predict a more potent version of the Py5OMe ligand?

Improving the catalytic performance of WOCs by experimentally modifying its ligands is a challenging and time consuming task. Having studied the Py5 based WOCs in detail, we asked ourselves whether we could predict a catalytically more active version of the ligand by altering the electronics and sterics of the dangling base.

Since altering the dissociation behavior of a chlorido substituent is not a straightforward process, we were curious whether other modifications of the catalyst could enhance its catalytic performance. To this end, we computationally modified the Py5OMe ligand in order to improve the water oxidation capability by introducing substituents at the dangling pyridine. The electronics of a pyridine and thereby the basicity of the nitrogen atom can easily be changed by placing functional groups in the *ortho*, *meta*, or *para* position. Such modifications are in principle also experimentally feasible, however the practicability has to be tested for each individual case. In a first step we checked whether such substitutions would affect the formation of the active catalyst, i.e. the $N_{Py}-Ru$ bond dissociation. For substituents in either the *meta* or *para* position we noticed a marginal difference, while *ortho* substituents in particular CF_3 led to a larger equilibrium bond distance and therefore weaker bond, as did to a smaller extent OMe. This suggests that *ortho* substituents affect both the sterics as well as the electronics of the system. Having established that entering the catalytic cycle should not be affected by the newly introduced substituents, we continued calculating the (absolute) barriers for the O–O bond formation. In general, substituents at the *ortho* position increased the barrier, in particular if the substituents were electron withdrawing and bulky at the same time such as CF_3. The size of substituents introduced in the *para* position has no compelling effect on the barrier, however electron donating substituents have the tendency to significantly lower the barrier — as illustrated by *p*-OMe, where the barrier went down to 12 kcal mol^{-1} compared to 16 kcal mol^{-1} for the unmodified pyridine.

Therefore this modification has the potential to enhance the catalytic performance of the catalyst drastically. To further reinforce this finding, we also compared the thermodynamics i.e. the change in free energy for the initial ET and PCET reactions

that are part of the proposed water oxidation mechanism. There the influence was again minimal, which was also expected since the pyridine is not directly affecting the electronics at the metal center. If an alteration of the overall thermodynamics is desired, the concept of an thermodynamical 'ideal' catalyst becomes handy.[70] This concept assumes that the water oxidation reaction is free of kinetic barriers, which implies that the minimal electrochemical force required to drive the reaction is equal to a quarter of the 4.92 eV necessary for water oxidation (see Equation 1.1). The comparison of the 'overall' thermodynamics of a catalyst or of potential WOC candidates obtained from *in silico* design to a thermodynamically 'ideal' catalyst can serve as a design criterion to identify the promising candidates as shown for various homogeneous catalysts.[67,131,133,175]

We analyzed various properties such as the spin density of the Ru center and the oxo-ligand, as well as the LUMO-energy of the AR, the TS and the associated product (AP) of the O−O bond formation as possible descriptors for the WNA reaction. Thereby, we found an indirect proportionality between the spin density at the oxo-ligand and barrier height. This suggests that a high spin density is consistent with larger electrophilicity of the oxo-ligand (see Chapter 5).[174]

This study shows that on the one hand the system has to be understood very well in order to rationally design and improve an existing system. On the other hand it appears that tuning of certain properties might be more straightforward than others, e.g. whereby one often faces the challenge not to deteriorate one aspect when improving another one. Durand *et al.* published a review where possible descriptors for the rational ligand catalyst design are presented.[176]

Applying Explicit Solvation

As showcased in the previous sections, applying a rather common simulation protocol including implicit solvation can provide insights into catalytic cycle, the behavior of ligands, and in some cases even allowed to propose possible improvements to the ligand-framework in order to enhance the catalytic performance. However, there is no doubt that for some properties, such as p$K_a$values, the inclusion of explicit solvation is of great importance.

Simulating pK_a Values by AIMD Simulations

Research Aim - Can we predict reliable pK_a values of transition metal complexes by AIMD?

Throughout this work the necessity to treat the solvent at the same level of theory as the solute became more and more obvious. Since transition metal base complexes are rarely studied employing AIMD the question arose which simulation protocol would be suitable to determine pK_a values of such complexes systems.

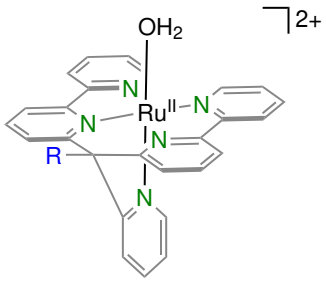
In section 1.2.5 we gave a brief overview of possible simulation protocols to compute p$K_a$values. In the following, we will focus our attention on the application of the *Bluemoon* methodology to determine the p$K_a$value of aqua ligands. For the full-length story describing the benchmarking of simulation protocols to determine

pK_a values see Chapter 6. In principle the simulation of p$K_a$values within the *Blue-moon* ensemble is a straight forward process. After setting-up a simulation cell that contains the protonated acid (AH) as well as several solvent molecules, all there is to do is to choose a CV that describes the deprotonation. Usually the A−H distance does a decent job. Now, several independent simulations where the constraint is continuously elongated have to be run in a consecutive manner. In the limit of a large enough constraint the force acting on the constraint will approach unity, indicating that the A^- state is reached. But is it? There is no way to tell by only considering the A−H distance, as the deprotonated acid could be reprotonated by different protons that originate form the solvent or oder acidic protons of the acid. For this reason Sprik proposed to use the coordination number (CN) - a CN is the generalization of a bond distance, where instead of 1:1 distance, a 1:N distance is monitored - of the acid instead of a distance constraint.[177] Nonetheless, it turned out that even when the simple distance constraint was used, often a good agreement with experiments could be achieved.[95,178,179]

As for the choice of the constraint, there is not a single protocol on how to relate the free energy obtained by the thermodynamic integration scheme into a p$K_a$value. Some authors argue that the obtained free energy directly corresponds to the deprotonation free energy, which is related to the dissociation constant according to Equation 6.4, we refer to this approach as 'absolute' pK_a.[90,178,180] Others prefer to calculate the dissociation constant according to Chandler's derivations using the radial distribution function of the protons with respect to the acid, weighted by the calculated free energy (see Equation 6.5).[181] By defining the dissociation constant in terms a radial distribution function, a cut-off radius for the A−H distance needs to be specified. This allows to distinguish between the protonated and deprotonated state. Again, different approaches exist, and while some authors determine the cut-off value by referencing the calculated p$K_w$value of H_2O to the experimental value (see Equation 6.8), in the later this approach is referred to as 'probabilistic' pK_a,[96,97,182–184] others prefer to calculate relative p$K_a$values according to Equation 6.6, termed by us as 'relative' pK_a.[179,185]

As there is no direct comparison among the different approaches in literature, we decided to evaluate the performance of those three post-processing methods for a series of small molecules as well as for different oxidation states of the $[Ru^{II}Py5R(H_2O)]^{2+}$ catalyst (see Figure 1.15). Note the p$K_a$values were evaluated for a catalytically inactive species where the chlorido ligand has been substituted by a solvent molecule.

In order to facilitate the comparison with protocols reported in literature we applied the simple A−H distance constraint. With our limited test-set we found that all three methods were able to qualitatively reproduce experimental p$K_a$values. The largest deviations were found in case of the 'relative' protocol. However, as it requires exactly the same simulations as the 'probabilistic' protocol there seems to be no necessity to use said approach. Overall the p$K_a$values obtained by the 'absolute' protocol appear to be slightly underestimated while the ones obtained by the 'probabilistic' protocol are slightly overestimated. This is in particular true for acidic compounds such as HCOOH and $[Ru^{III}Py5R(H_2O)]^{3+}$. Those deviations deserve more attention, however the large computational cost make the simulation of large benchmark libraries unfeasible. We further were able to show that the p$K_a$is not dramatically affected size of the simulation cell, which further strengthened our



$[Ru(II)Py_5R(OH_2)]^{2+}$, where R = Me or OMe

Figure 1.15: Lewis structure of $[Ru^{II}Py5R(H_2O)]^{2+}$, where R = Me or OMe. Reproduced from Schilling *et al.*[174]

confidence in the simulations.

With this study we haven shown that p$K_a$values of transition metal complexes can be simulated successfully within the *Bluemoon* ensemble. Further, the direct comparison between the different post-processing methods can serve as a guideline for further studies.

Simulating the O−O Bond Formation by AIMD Simulations

Research Aim - How do solvent-solute interactions influence the O−O bond formation?

The sophisticated description of the solvent and solute including finite temperature effects by AIMD simulations allows us to assess the quality of the mechanistic predictions done employing implicit solvation by

Due to the explicit description of the solvent and solute, AIMD can assess the quality of the mechanistic predictions proposed by simpler solvation models. This allowed us to answer the following questions: (1) How do the reaction mechanisms obtained by different methods i.e. implicit solvation, *Bluemoon* and MetaD differ in terms of their energetics and the structures of the reactant, TS and product. (2) Which alternative reaction mechanisms for the O−O bond formation exist that involve the solvent but not the dangling base.

In the previous sections we have introduced the Ru-based WOC $[Ru^{V}Py5MeO(O)Cl]^{2+}$ and presented TS structures for the O−O bond formation by a WNA mechanism (see Figure 1.12). Localization of the those extrema on the potential energy surfaces, was only possible if at least one additional explicit solvent molecule was included in the model system. This implies that solute-solvent interactions are crucial for the O−O bond formation. The reaction is therefore predestined to be modeled using AIMD simulations, whereby the large activation barrier (16 kcal mol^{-1}) as compared to k_BT (0.593 kcal mol^{-1}) make enhanced sampling methods such as *Bluemoon* or MetaD necessary. Both depend on the definition of a CV that describes the chemical transformation of interest (see section 2.4). For the full-length story of the O−O bond formation modeled by AIMD see Chapter 7. An obvious choice of a CV to describe the O−O bond

formation by either a base-assisted or base-independent WNA is the O−O distance (see Figure 1.16). We modeled both mechanisms within the *Bluemoon* ensemble applying said CV. Such a comparison would have been very challenging within the methodology applied in our previous studies due to the fact that the formed hydronium is very difficult to stabilize by a limited number of explicit solvent molecules.

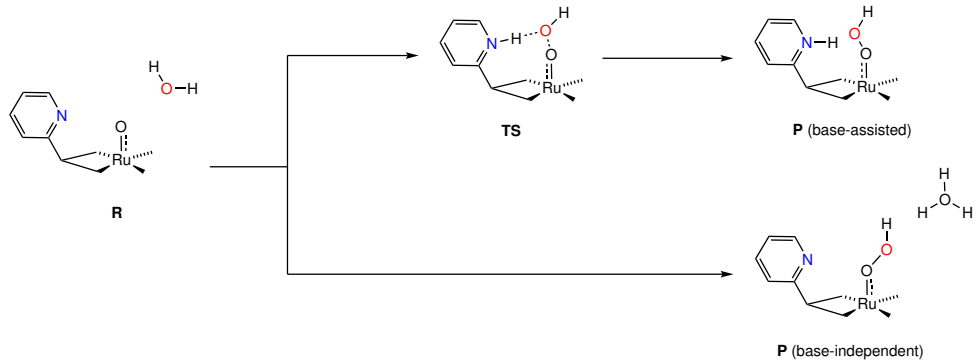


Figure 1.16: Reaction scheme for the base-assisted and base-independent WNA mechanism for the O−O bond formation catalyzed by $[Ru^{V}Py5MeO(O)Cl]^{2+}$ (*unpublished figure*).

The obtained activation barriers clearly favored a base-assisted mechanism and thereby verified our previous mechanistic proposals. However, an in-depth examination of the base-assisted reaction path revealed that the O−O bond distance was insufficient to fully describe said mechanism. This is showcased by the fact that the proton was found to be either located at the nucleophile or at the base. The nonexistence of intermediate states suggested that the proton transfer was not captured properly.

This problem was overcome by employing a difference of distances constraint - in this case d(O−O)-d(O−H). While, including the O−H distance in the CV resulted in an uniform sampling of both the O−O distance as well as the O−H distance another problem arose. Namely, the indistinguishability of the two protons of the nucleophile at large distance where neither of them interacts with the base. This is in particular problematic if the proton, that is not involved in the CV, in the course of the of the O−O bond formation engages in hydrogen bonding with the base. In order to overcome those challenges a methodology less restrictive in terms of the applied CV than *Bluemoon* was required. Here MetaD where multiple independent CVs can be used promised to be a reasonable alternative.

Over the course of several simulations, a suitable set of CVs was identified. One of the CVs was chosen to keep track track of both the protonation state of the nucleophile and the base, while the other describes the O−O distance. Applying the WT-MetaD formalism,[47] where the height of the spawned Gaussian is continuously decreased over the course of the simulation. A well known problem in the context of MetaD is the assessment of convergence. In particular in the context of AIMD simulations where the available computational resources tends to become a limiting factor. Here the multiple reoccurrence of the O−O bond formation is taken as a strong indication of convergence. This is supported by the low statistical errors of

the local extrema of interest i.e. the R, the TS, and the P. The MEPs connecting those extrema were identified by a shortest-path algorithm (see Figure 1.17).[186]

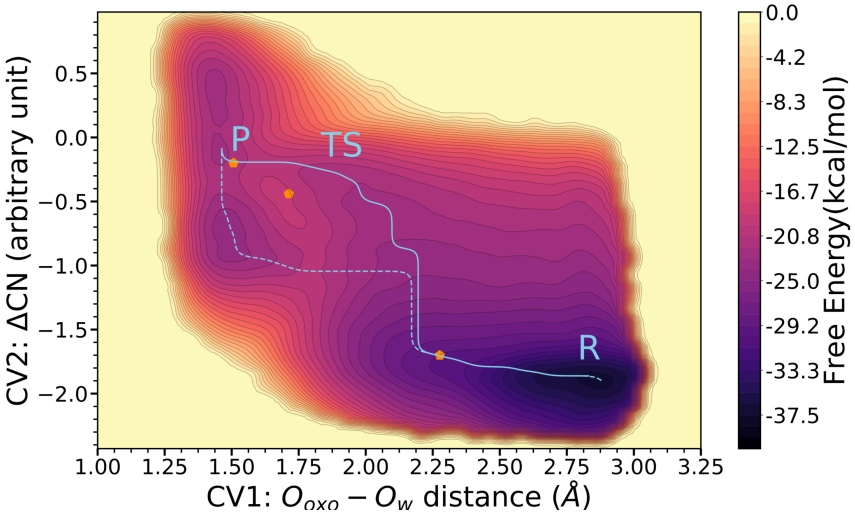


Figure 1.17: Free energy surface of the O−O bond formation catalyzed by $[Ru^{V}Py5Me(O)Cl]^{2+}$ reconstructed from WT-MetaD simulations. The MEPs of the base-assisted (solid line) and base-independent (dashed line) connecting the R and P are shown. The orange pentagons represent the AR, TS and P obtained by the geometry optimization based DFT approach used in Chapter 5 (*unpublished figure*).

The apparent co-existence of two MEPs, one describing the base-assisted pathway and one the base-independent mechanism showcases a major problem of any dimensionality reduction procedure, the loss of information. By projecting the free energy surface along an additional dimension, in this case the protonation state of the base, it becomes obvious that no transition along the base-independent mechanism was observed during the simulation (see Figure 7.12). In agreement with both the static-DFT and the *Bluemoon* simulations the base-assisted pathway was identified as the energetically most favorable. When comparing the MEP and the extrema of interest localized by the different methods no significant structural differences can be found. However, in case of MetaD the P state turned out to be much more versatile. Besides the expected hydroperoxo species, also its deprotonated form the peroxo species as well as hydroperoxo species where the base has been deprotonated were explored. This implied that both protons of the nucleophile become acidic in the course of the O−O bond formation. While the proton transfer to the base was found to be favorable in case of the O−O bond formation, an increased tendency of formed hydroperoxo species to release its proton to the solvent is found. Supporting the general assumption that the hydroperoxo species is usually a highly elusive intermediate that quickly undergoes further electron and proton transfer reactions in order to release molecular oxygen.

In this study we have demonstrated how the choice of an appropriate model system and methodology can lead to a more fundamental understanding of the catalytic cycle. Explicit solvation in particular allows to explore intermediates and reaction pathways that are inaccessible in less elaborated simulation protocols.

Chapter 2

Applied Computational Methods

The aim of this chapter is to give a brief introduction of the different methods used in this work. In this sense, we neither attempt to give complete review of all the methods used in this field nor to give a detailed mathematical derivation. We refer the interested reader the corresponding original literature.

In quantum chemistry a system is described by a wavefunction Ψ that encodes for all information. As Ψ is *a priori* unknown, the main goal of computational chemistry is to derive a good enough approximation in order to evaluate the properties of the system.

In order to do so the famous Schrödinger equation has to be solved. This can only be done exact for very small systems. If only ground state properties of a system containing non-heavy atoms are of interest, it is often sufficient to do so only for the non-relativistic time-independent case:

$$\hat{\mathcal{H}}|\Psi\rangle = E|\Psi\rangle, \tag{2.1}$$

where $\hat{\mathcal{H}}$ is the Hamiltonian of the system, Ψ the wavefunction and E the energy.

Due to the variational principle the optimization of approximate wavefunctions gives an upper bound for the interacting ground state energy. The 'best', approximative wavefunction Ψ can therefore be obtained by minimizing the energy with respect to the trial wavefunction Ψ'.

There are various approaches on how to solve the Schrödinger equation approximately, among them are wavefunction based approaches such as Hartree–Fock (HF), configuration-interaction (CI), coupled cluster (CC), but there is also density functional theory (DFT). Nowadays, DFT is the work-horse of computational chemistry due to its favorable compromise between computational efficiency and performance. Because of its extensive usage throughout this work it will be briefly introduced in the following.

2.1 Density Functional Theory

The main idea behind DFT is a dimensionality reduction of the problem. In general, a wave function of an N electron system depends on 4N variables $\Psi(\mathbf{r_1}, \mathbf{r_2}, ..., \mathbf{r_N})$. Those are the three spatial coordinates and the spin of each electron ($\mathbf{r_i}$, where i

the index of the electron). On the other hand, the electron density of a system $\rho(\mathbf{r})$, where $\mathbf{r}$ is its positional vector, depends only three spatial coordinates and is independent of the number of electrons. Several theorems justify this ansatz:

Hohenberg–Kohn Theorems

Within the Born–Oppenheimer approximation,[187] the electronic Hamiltonian describing the electrons in the external field of the nuclei can be written as the sum of the kinetic energy operator $\hat{\mathrm{T}}$ for the electrons, the electron–electron interaction operator $\hat{\mathrm{V}}_{\mathrm{ee}}$, and the external potential operator $\hat{\mathrm{V}}_{\mathrm{ext}}$:

$$\hat{\mathcal{H}} = \hat{\mathrm{T}}_{\mathrm{e}} + \hat{\mathrm{V}}_{\mathrm{ee}} + \hat{\mathrm{V}}_{\mathrm{ext}}. \tag{2.2}$$

According to the first Hohenberg–Kohn theorem the ground state external potential, which is caused by the charge of the nuclei, is up to a constant uniquely related to the ground state total electron density ρ_0.[188] The expectation value of the external potential operator can be expressed as:

$$\left\langle \Psi_0 \left| \hat{\mathrm{V}}_{\mathrm{ext}} \right| \Psi_0 \right\rangle = \int \rho_0(\mathbf{r}) \cdot V_{\mathrm{ext}}(\mathbf{r}) d\mathbf{r}, \tag{2.3}$$

where Ψ_0 is the ground state wave function.

The ground state energy of the many-body system can therefore be written as a functional of ρ_0:

$$E\left[\rho_0\right] = T\left[\rho_0\right] + V_{ee}\left[\rho_0\right] + \int \rho_0(\mathbf{r}) \cdot V_{\mathrm{ext}}(\mathbf{r}) d\mathbf{r}. \tag{2.4}$$

The second Hohenberg–Kohn theorem states that the variational principle can be applied in order to approximate the energy of the electronic ground state $E\left[\widetilde{\rho}_0\right]$ and find the exact ground state density ρ_0:[188]

$$E\left[\widetilde{\rho}_0\right] \geq E\left[\rho_0\right], \tag{2.5}$$

where $\widetilde{\rho}_0$ is a trial electron density.

Kohn–Sham Density Functional Theory

Unfortunately, without knowledge of the 'exact' functional form of the kinetic energy $T\left[\rho\right]$ and the electron–electron interaction $V_{ee}\left[\rho\right]$ (see Equation 2.4), attempting to find ρ that minimizes the energy to a reasonable accuracy is beyond hope.

To overcome this problem, Kohn and Sham showed that there is potential, whose ground state electron density of the non-interacting systems is the same as the one of the interacting system.[189] For the non-interacting system $\hat{\mathcal{H}}$ (in atomic units) can be expressed as:

$$\hat{\mathcal{H}} = \hat{\mathrm{T}} + \hat{\mathrm{V}}_{\mathrm{eff}} = -\frac{1}{2}\sum_{i=1}^{N} \nabla_i^2 + \sum_{i=1}^{N} V_{\mathrm{eff}}(\mathbf{r_i}), \tag{2.6}$$

where N is the number of electrons, and $i \in N$ their index. Note that all terms of $\hat{\mathcal{H}}$ are one electron operators. The electron density is related to the occupied one electron functions (orbitals) φ_i of the non-interacting quasi-electrons as:

$$\rho(\mathbf{r}) = \sum_{i=1}^{N^{occ.}} |\varphi_i(\mathbf{r})|^2 , \tag{2.7}$$

where $N^{occ.}$ is the number of occupied orbitals and $i \in N^{occ.}$ the index of the orbital. Now, Equation 2.4 can be reformulated as:

$$E[\rho] = T_s[\{\varphi_i\}] + J[\rho] + E_{XC}[\rho] + E_V[\rho] , \tag{2.8}$$

where the kinetic energy of the non-interacting system functional is:

$$T_s[\{\varphi_i\}] = \sum_{i=1}^{N} \left\langle \varphi_i(\mathbf{r}) \left| -\frac{1}{2}\nabla_i^2 \right| \varphi_i(\mathbf{r}) \right\rangle . \tag{2.9}$$

The coulomb interaction functional is:

$$J[\rho] = \frac{1}{2} \iint \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r} d\mathbf{r}' \tag{2.10}$$

$$= \frac{1}{2} \sum_{i}^{N} \sum_{j>i}^{N} \int \int |\varphi_i(\mathbf{r})|^2 \frac{1}{|\mathbf{r} - \mathbf{r}'|} |\varphi_j(\mathbf{r}')|^2 d\mathbf{r} d\mathbf{r}', \tag{2.11}$$

where i and $j \in N$ are the indexes of the electrons. Further, the nucleus-electron interaction functional is:

$$E_V[\rho] = \int \rho(\mathbf{r}) \cdot V_{\text{ext}}(\mathbf{r}) d\mathbf{r} \tag{2.12}$$

$$= -\sum_{i}^{N} \int \sum_{A}^{M} \frac{Z_A}{|\mathbf{r_i} - \mathbf{R_A}|} |\varphi_i(\mathbf{r})|^2 d\mathbf{r}, \tag{2.13}$$

where M the number atoms in the system, $A \in M$ is the index of the atom, Z_A its the core charge.

Finally, the *exchange–correlation* functional $E_{XC}[\rho]$ accounts for the exchange (X), correlation (C) and the difference between the exact kinetic energy and the one of the non-interacting system. The exact form of this functional is unknown, therefore it has to be approximated.

Local Density Approximation

A first approximation of $E_{XC}[\rho]$ is the so called local density approximation (LDA), the foundation of which is the uniform electron gas, where electrons are assumed to move on a positive background charge distribution:[190]

$$E_{XC}^{LDA}[\rho] = \int \rho(\mathbf{r}) \varepsilon_{XC}^{LDA}[\rho(\mathbf{r})] d\mathbf{r}, \tag{2.14}$$

where $\varepsilon_{XC}^{LDA}[\rho(\mathbf{r})]$ is the functional derivative of $E_{XC}^{LDA}[\rho(\mathbf{r})]$. It can be subdivided into its exchange and correlation component:

$$\varepsilon_{XC}^{LDA}[\rho(\mathbf{r})] = \varepsilon_{X}^{LDA}[\rho(\mathbf{r})] + \varepsilon_{C}^{LDA}[\rho(\mathbf{r})]. \tag{2.15}$$

The exchange component is known in the case of LDA:

$$\varepsilon_{X}^{LDA}[\rho(\mathbf{r})] = -\frac{3}{4}\Big(\frac{3\rho(\mathbf{r})}{\pi}\Big)^{\frac{1}{3}}. \tag{2.16}$$

For the correlation part no analytical form is known, consequentially it is usually obtained by fitting to a higher level of theory.[191]

Generalized Gradient Approximation

Within the LDA ε_{XC}^{LDA} depends only on density at specific points ($\mathbf{r}$), this approximation turned out to be inadequately describe non-homogeneous systems such as molecules. To this extent, the generalized gradient approximation (GGA), where the gradient of the density ($\nabla\rho(\mathbf{r})$) is also taken into account, turned out to be a significant improvement over LDA.[192,193]

$$E_{XC}^{GGA}[\rho] = \int \rho(\mathbf{r})\varepsilon_{XC}^{GGA}[\rho(\mathbf{r}), \nabla\rho(\mathbf{r})]\, d\mathbf{r} \tag{2.17}$$

Exchange–Correlation Functional Zoo

The need to included even more information of the density into the exchange-correlation functional led to the development of meta-GGA functionals that depend also on the Laplacian of the orbitals. Then there are hybrid functionals, where the exact exchange contributions from HF are added. There is also the combination of both i.e. hybrid-meta-GGA functionals. Besides that there are also range-separated hybrid and local hybrids, where the electron-electron interactions are treated position-dependent, and combinations thereof.[194] The huge number of exchange–correlation functionals all of which have their slightly different formulation and intended field of application make the choice of the 'best' functional for the task at hand tedious.[195] There are attempts to thoroughly benchmark all functionals in the 'zoo' for specific applications in order to allow for a fair comparison.[196] In principle, choosing the 'right' functional is a multi-variable optimization that is hardly ever feasible. That is why some functionals such as PBE[192] and B3LYP[197,198] are over-proportionally used.[195] Even though they are probably not optimally suited for all the tasks they were used for, but their popularity at least allows for a qualitative comparison of different contributions, which would be hardly possible if all known functionals would be used with a similar frequency.

Solving the Kohn–Sham Equations

Having defined all terms of Equation 2.8, solving of Kohn-Sham (KS) equations becomes possible.

$$\left[-\frac{1}{2}\nabla^2 + V_{\text{eff}}(\mathbf{r})\right]\psi_i(\mathbf{r}) = \varepsilon_i\psi_i(\mathbf{r}), \tag{2.18}$$

where ψ_i is a KS orbital and ε_i its corresponding the energy. V_{eff} contains the contributions from J, E_v, and E_{XC} (see Equations 2.10, 2.12 and 2.18).

The KS equation are solved within the self-consistent field (SCF) approach, whereby the coefficients of the basis functions that define the KS-orbitals are varied until the energy convergences to a user-defined accuracy. Optimization tasks such the one at hand are perfectly suited for the application on computers - and are some sense the 'essence' of computational chemistry.

2.1.1 Dispersion

When it comes to the accurate description of chemical systems KS-DFT does in general a decent job. However, one of the limitations of the functional approximation used in KS-DFT is the absence of long-range interactions i.e. London-forces. Those are in particular crucial for intermolecular interactions. Probably the most popular semiempirical correction scheme to overcome this problem was developed by Grimme and co-workers, it us usually referred to as 'D3'[199–201].

It is based on two-body interaction potentials, which are summed up and added to the KS-energy.

$$E_{disp}^{D3} = -\sum_{A,B}\sum_{n=6,8} s_n \frac{C_n^{A,B}}{r_{A,B}^n + f_n(r_{A,B})} \tag{2.19}$$

where A and B are the indexes of two interacting atoms, s_n is a functional depended scaling factor, $C_n^{A,B}$ are atom-in-molecules coefficients calculated by TD-DFT, f_n is a damping function to avoid double-counting in the short-range regime. Nowadays, the D3 correction or one of its alternatives are routinely included in all KS-DFT simulations.

There are also other correction schemes known to account for long-range interactions for example: A family of nonlocal van der Waals functionals named after Vydrow Van Voorhis (VV)[202] or the density dependent dispersion correction by Tkatchenko and Scheffler (TS-vdW).[203] However, neither of them was used in the current work.

2.2 Solvation Models

Many chemical reactions, in particular water oxidation discussed here, take place in the condensed phase. In Section 1.2.4 several approaches on how to treat solute–solvent interactions are briefly introduced. Thereby the classification of those various approaches into: explicit and implicit solvation, depending on the description of the solvent as well as static or dynamic depending on the simulation methodology is discussed.

From a methodological point of view, static explicit solvation is the same as applying KS-DFT, discussed in the previous section, to a system containing multiple molecules. On the other hand, dynamic explicit solvation requires, as the term implies, a dynamic treatment of the solvent. The latter is usually achieved by molecular dynamics, the foundation of which is discussed in 2.3.

A fundamentally different approach is the concept of implicit solvation. For an exhaustive discussion of its physical background and details on the most common implementations we refer to the work by Tomasi.[62]

Here a brief introduction for the commonly applied COSMO solvation model is given.[59,204,205] In order to describe the solute–solvent interaction it is assumed that the solute generates a cavity within the dielectric continuum representing the solvent. To allow for a homogeneous treatment of the electrostatic screening of the solute by the solvent, the surface of the cavity has to be subdivided into m segments. The charge distribution of the solute causes electrostatic potentials acting on each segments of the cavity, that can be represented as the vector $\mathbf{\Phi} = (\Phi_i, ..., \Phi_m)$. The latter cause a polarization of the continuum, the response of which is interpreted as screening charges $\mathbf{q}$. The interactions of those screening charges is described by a Coulomb interaction matrix $\mathbf{U}$, the setup of which prevents the self-interaction of screening charges.

The key assumption of COSMO is that the screening charges in a conducting continuum (i.e $\varepsilon = \infty$) $\mathbf{q}^*$ and the electrostatic potential arising from the solute $\mathbf{\Phi}$ cancel each other.

$$0 = \mathbf{\Phi} + \mathbf{U}\mathbf{q}^*. \tag{2.20}$$

Further, for a continuum with a finite permittivity (ε) the screening charges might be approximated as:

$$\mathbf{q} = f(\varepsilon)\mathbf{q}^* \tag{2.21}$$

with

$$f(\varepsilon) = \frac{\varepsilon - 1}{\varepsilon + x} \tag{2.22}$$

where x is an empirical parameter - usually set to $\frac{1}{2}$ for neutral molecules.

The dielectric energy i.e. the interaction between the solute and solvent is therefore:

$$E_{diel} = \frac{1}{2} f(\varepsilon)\mathbf{\Phi}\mathbf{U}^{-\mathbf{1}}\mathbf{\Phi} \tag{2.23}$$

the term $\frac{1}{2}$ arises from the fact that half of the interaction energy is required to polarize the continuum.

The dependence of E_{diel} on the charge distribution of the solute makes it necessary to include the calculation it into the self-consistent-field optimization of the wavefunction or the KS-equations (in the case of KS-DFT).

Like all continuum solvent models (CSMs), COSMO depends on an empirical parameterization, which involves atomic radii used to build the cavity, the empirical scaling factor in Equation 2.22 as well as the permittivity describing the solvent. The empirical parameterization can lead to confusions regarding which contributions to,

for example, the solvation free energy are included in the result of a CSM calculation.[65] This also manifests in some disagreement on how certain quantities such as pK_a values have to be calculated.[79]

Over the years, the original COSMO methodology has been continuously improved and extended which led to COSMO-RS.[59] The main idea of which is to use the COSMO screening charge density of both the solvent and solute to calculate the chemical potential of the system according to statistical thermodynamics. This method allows to accurately describe solvent mixtures and ionic liquids, which are beyond the capability of a simple CSM. In the Turbomole program package an improved version of the COSOM solvation model termed DCOSMO-RS has been implemented.[59,206] DCOSMO-RS is a self-consisted implementation of COSMO-RS that requires a σ potential obtained from COSMO-RS.

2.3 Molecular Dynamics

MD simulations are used when a single atomistic configuration and its properties cannot sufficiently represent the properties of the system anymore. This is in principle the case for all systems that involve either multiple or highly flexible molecules. In such cases a statistical interpretation is required. There are different methodologies known for generating statistical ensembles. On one hand there are Monte-Carlo schemes where conformations are sampled by random trial-moves that are either accepted or refused based on some predefined rules such as for example the total energy. On the other hand there is the molecular dynamics methodology, where the statistical ensemble is obtained by propagating the system in time. The fundamental assumption behind this approach is the ergodicity hypothesis, which states that for 'long enough' time series the time average converges to the ensemble average:

$$\lim_{t\to\infty} \frac{1}{t} \int_0^t X(t)dt = \int_{-\infty}^{\infty} x \cdot \rho_X(x)dx \tag{2.24}$$

where $X(t)$ is the time series of the process of interest, x a micro-state of this process, t the time, and $\rho_X(x)$ probability density function of $X(t)$.

A prerequisite to propagate a system in time is the *a priori* knowledge of the interactions between all particles of the many-body system. Within the BO approximation the nuclear and electronic dynamics are decoupled (BO-MD). This justifies a description of the nuclear motion according to Newton's equation of motion. The contributions of the electrons describing the particle-particle interaction is this classical picture, approximated as a potential which is known as force-fields. In their simplest form they consist two-body interactions such as Coulomb potentials for charged particles and Lennard-Jonnes potentials that account approximatively for the Pauli repulsion at short ranges and van der Waals forces at long ranges.

Because Newton's equation of motion cannot be solved analytically for a many-body system, a finite difference scheme is applied. To do so one assumes that the forces acting on the particles remain constant within a time-step Δt. From the particle positions $\mathbf{R}$ and the forces acting on them, their accelerations $\mathbf{A}$ and velocities $\mathbf{V}$ can be calculated. With those at hand, the position of particles at the next time step $\mathbf{R}(t + \Delta t)$ can be obtained.

An algorithm to do so is the called position-Verlet[207] and takes advantage of a Taylor expansion in order to define the position of the I-th particle at the time $t + \Delta t$ ($\mathbf{R_I}(t + \Delta t)$):

$$\mathbf{R_I}(t + \Delta t) = \mathbf{R_I}(t) + \mathbf{V_I}(t)\Delta t + \mathbf{A_I}\Delta t^2 + \frac{\Delta t^3}{3!}\frac{\partial^3 \mathbf{R_I}}{\partial t^3} + \mathcal{O}(\Delta t^4) \tag{2.25}$$

where $\mathcal{O}$ is the error of a truncated series.

It can be shown that the velocity of a particle is defined as:

$$\mathbf{V_I}(t) = \frac{\mathbf{R_I}(t + \Delta t) - \mathbf{R_I}(t - \Delta t)}{2\Delta t} + \mathcal{O}(\Delta t^2). \tag{2.26}$$

This is problematic since in order to calculated velocities at $t = 0$ the positions of the previous step $\mathbf{R}(t - \Delta t)$ have to be known. In practice therefore a reformulation of the algorithm known as velocity-Verlet is used.[208]

Velocity Verlet

Expressing Equation 2.25 at time $t - \Delta t$ i.e. ($\mathbf{R_I}(t - \Delta t)$) and evaluating it at $t + \Delta t$ leads to:

$$\mathbf{R_I}(t) = \mathbf{R_I}(t + \Delta t) - \mathbf{V_I}(t + \Delta t)\Delta t + \frac{\mathbf{A_I}}{2}\Delta t^2 - \frac{\Delta t^3}{3!}\frac{\partial^3 \mathbf{R_I}}{\partial t^3} + \mathcal{O}(\Delta t^4) \tag{2.27}$$

to which Equation 2.25 can be added, which results in:

$$\mathbf{V_I}(t + \Delta t) = \mathbf{V_I}(t) + \frac{\Delta t}{2}\Big(\mathbf{A_I}(t) + \mathbf{A_I}(t + \Delta t)\Big) + \mathcal{O}(\Delta t^4). \tag{2.28}$$

An alternative to the Verlet algorithms is the so called 'leapfrog' algorithm by by Van Gunsteren and Berendsen, which is similar in terms of performance and accuracy.[209]

In a typical simulation the number of particles (N) remains the same throughout the simulation, and so does the simulation cell size i.e. the volume (V). In this case, Newton's equations of motion assure convergence of the total energy (E) which is the sum of the potential and kinetic energy of all the particles in the system. In statistical mechanics, a system where N, V and E are kept constant (NVE) is referred to as micro-canonical ensemble. The total energy is an inconvenient observable, therefore often either the temperature (T) or the pressure (P) are kept constant as such as in the NVT (canonical ensemble) or NPT ensemble (isothermal-isobaric ensemble). Constant temperature is usually enforced by coupling the system to an external heat-bath that allows particles to exchange kinetic energy. For this purpose commonly the Nosé-Hoover thermostat is used which is based of an extended Lagrangian formalism to restrain the temperature of the system.[210,211]

2.3.1 Ab Initio Molecular Dynamics

The application of classical molecular dynamics is often limited by the required *a priori* knowledge of the potential i.e. force-field describing the particle-particle

interactions in system. The construction of the a force-field as well as its parameterization dictate the outcome of the simulation. Further, most classical force fields are not able to appropriately describe the formation or breakage of chemical bonds.

To this extent *ab initio* methods offer to be out of the box solution. Those methods rely on the calculation of the electronic structure for each sampled atomic configuration. From the total energy and the position of the nuclei, then the forces required to propagate the system can be obtained. As in the case of classical MD, the BO approximation justifies the application of Newton's equation of motion to propagate the system.

The costly minimization of the electronic wavefunction at each step of the trajectory led to the development of the Car-Parinello MD (CPMD) formalism.[212] Wherein the electronic degrees of freedom are included in the propagation step, thereby reducing the computational cost for each step by avoiding the complete minimization of the electronic wavefunction. Nowadays however, due to the increase in computational power, the method lost some its initial attractiveness.

Independent of the method used to obtain forces a typical AIMD scheme looks as follows:

- calculate the total energy of the system $E(\mathbf{R}(t))$
- calculate forces
- (optional) apply thermo- or barostat
- propagate nuclei by the forces to obtain their positions at time $t + \Delta t$ ($\mathbf{R}(t + \Delta t)$)
- repeat until a certain property of the system is 'converged' i.e. a until the ensemble average of that property reached the desired accuracy

2.4 Enhanced Sampling Methods

In the context of molecular dynamics simulations, the time-scale on which the event of interest takes place is of uttermost importance. The event has to be observed in a statistically relevant manner in order to make reliable statments regarding its properties. Nowadays the limit in terms of simulation length for atomistic force-field simulations of reasonably large systems with several thousands of atoms is in the range of μs. For *ab initio* MD, on the other hand time-scales in the range of a few *ns* for systems containing several hundreds up to a few thousand atoms is the absolute limit.

However, those time-scales are usually far too short in order to investigate interesting real-life problems. For example enzymatic reactions, that are coupled to major conformational changes, often occur in time-scales of milliseconds to seconds. Those are out of reach by any kind of simulation for the time being.[213]

This limitation can to some extent be overcome by enhanced sampling techniques, where instead of exploring the full phase space only the subspace where the event of interest occurs is sampled. Such methods included but are not limited to umbrella sampling,[214] *Bluemoon*,[92] and MetaD[45]. For a more complete overview of those methods see Yang *et al.*[215]

Most enhanced sampling methods rely on so-called CVs, which are used to distinguish between different micro-states of the system. The most simple one are intra- or intermolecular distances and angles. However, in principle there exist no restrictions when it comes to define a CV, as long as it is a continuous differentiable function that maps a micro-state to a scalar. Even exotic properties such as the conformational entropy might be used.[216] Nonetheless, it is well known that the choice of a proper CV that describes the event interest is a huge challenge. There are attempts to automatically identify collective variables.[217] However, often the choice is still based on the chemical intuition of the person who runs the simulations.

2.4.1 Bluemoon

The *Bluemoon* methodology in principle is nothing else than a thermodynamic integration scheme, that is used to calculate the differences in free energy ΔF. Thereby a single CV ($\xi(\mathbf{R})$) is used to describe the event of interest. The value range of the CV relevant for the event is discretized, and for each value a constrained MD simulation has to be carried out, during which the force required to enforce the constraint is evaluated ($f_{\xi'}$). Integrating the average forces over the range of the CV results in the free energy (a detailed derivation of the methodology can be found in literature):[92–94]

$$\Delta F \;=\; -\int_{\xi_0}^{\xi_1} f_{\xi'} d\xi'. \tag{2.29}$$

The average force $f_{\xi'}$ is derived from the Lagrange multiplier λ according to

$$f_{\xi'} \;=\; \frac{\langle Z^{-1/2}[\lambda - k_\mathrm{B}TG]\rangle_{\xi'}}{\langle Z^{-1/2}\rangle_{\xi'}}, \tag{2.30}$$

where k_B is the Boltzmann constant, T the temperature, and Z and G are correction factors associated with the transformation from generalized to Cartesian coordinates. Z is also known as the 'Fixman mass-metric tensor correction' and is defined as[218]

$$Z = \sum_{i=1}^{M} \frac{1}{m_i} \left(\frac{\partial \xi}{\partial \mathbf{R_i}} \right)^2 \tag{2.31}$$

where M is the number of atoms, $i \in M$ is the atomic index, m_i is the atomic mass, and $\mathbf{R_i}$ it's positional vector. The G term, accounting linearly dependent terms is defined as:

$$G = \frac{1}{Z^2} \sum_{i=1}^{M} \sum_{j=1}^{M} \frac{1}{m_i} \frac{1}{m_j} \frac{\partial \xi}{\partial \mathbf{R_i}} \cdot \frac{\partial^2 \xi}{\partial \mathbf{R_i} \partial \mathbf{R_j}} \cdot \frac{\partial \xi}{\partial \mathbf{R_j}} \tag{2.32}$$

where i and $j \in M$ are the indexes the atoms.

In case the special case where the CV is chosen to be the distance between two nuclei i and j $\left(\xi = r_{ij} = \sqrt{(\mathbf{R_i} - \mathbf{R_j})^2}\right)$, Z and G simplify to constant values. Therefore, Equation 6.2 can be rewritten as $f_{\xi'} = \langle \lambda \rangle_{\xi'}$. In all other cases where the order

parameter involves multiple distances (for example $\xi = r_{ij} - r_{jk}$). The difference of two distances, Z and G have to be analytically derived and evaluated for each step of the simulation.

In this case the analytical expressions for Z and G are[219]

$$Z = \frac{1}{m_i} + \frac{1}{m_i} + \frac{1}{m_k}\Big(1 - \frac{\mathbf{R_{ij}}}{r_{ij}} \cdot \frac{\mathbf{R_{kj}}}{r_{kj}}\Big), \tag{2.33}$$

and G = 0.

Therefore Equation 6.2 simplifies to

$$f_{\xi'} = \frac{\langle Z^{-\frac{1}{2}} \lambda \rangle_{\xi'}}{\langle Z^{-\frac{1}{2}} \rangle_{\xi'}}. \tag{2.34}$$

As stated earlier the choice of the appropriate CV is crucial in the context of most enhance sampling methods. In the case of the *Bluemoon* ensemble this is of uttermost importance since a highly complex 3N dimensional free energy surface is reduced to a single dimension. However, if a suitable CV has been identified the accuracy of the free energy can be increased using a finer grid for the thermodynamic integration.

2.4.2 Metadynamics

For more complex systems where a single CV is unable to properly describe the event of interest, methods such as MetaD have been developed. The basic principle of which is the introduction of a history dependent bias potential $V(\xi(x), t)$. This forces the system to explore the phase space of interest by penalizing it for the repeated visit of micro-states. The bias potential itself is usually composed of Gaussian functions with a predefined height and width:

$$V(\xi, t) = \int_0^t dt' \, \frac{W}{\tau} \, exp\left(- \sum_{i=1}^{d} \frac{(\xi_i - \xi_i(t'))^2}{2\sigma_i^2} \right) \tag{2.35}$$

where ξ is the value of the CV at time t, $\xi(t')$ is the position of the Gaussian at time t', d are the number of CVs, W the height, τ the deposition stride, and σ_i the width of the Gaussian of the i-th CV. It can be shown that in the limit of infinite time the bias potential converges to the free energy:

$$\lim_{t \to \infty} V(\xi, t) = -F(\xi) \tag{2.36}$$

The convergence of MetaD in its classical formulation is very difficult to assess since the constant increase of the bias potential enforces the exploration of high energy micro-states. This problem can be limited by the so called WT-MetaD formalism. In WT-MetaD, the bias deposition rate decreases over the course of the simulation. This ensures convergence of the free energy to a constant value. In practice, the reduction of the bias deposition rate is achieved by rescaling the Gaussian height according to:

$$W(\xi t) = W_0 \, exp\left(- \frac{V(\xi, t)}{k_B \Delta T} \right) \tag{2.37}$$

where k_B is the Boltzmann constant and ΔT a temperature difference. Instead of specifying ΔT it is common to define the so-called bias factor γ:

$$\gamma = \frac{\Delta T}{T + \Delta T} \tag{2.38}$$

WT-MetaD is not the only extension to the MetaD methodology (there are, for example, replica-exchange MetaD, bias-exchange MetaD or path-MetaD). For a detailed derivation of the method we refer to the original literature and corresponding review articles.[45–47,220–223]

2.5 General Computational Settings

The choice of optimal computational settings is always dependent on the task at hand. Here we give a general overview for a more detailed description and a justification of the specific settings used see the corresponding 'computational settings' section in the respective chapters.

2.5.1 Turbomole

Static-DFT calculations were usually carried out using the Turbomole software package (version 7.01).[206] In general, geometry optimizations and frequency analysis were performed employing (unrestricted) KS-DFT with the BP86 exchange-correlation functional,[224,225] Ahlrich's triple-ζ valence polarization basis set (def2-TZVP)[226] (with the corresponding effective core potential for transition metals such as Co and Ru and lanthanides such as Er and Tm) and Grimme's D3 dispersion correction[199] (short hand notation: BP86-D3/def2-TZVP). For each optimized structure normal mode analysis using either AOFORCE or NUMFORCE[227,228] was performed to confirm their nature as minimum on the potential energy surface.

Single-point electronic energies were usually obtained employing the hybrid-GGA functional B3LYP[197,198] and the def2-TZVP basis set. In addition Grimme's D3 dispersion correction and the COSMO solvation model[204] (dielectric constant ε = 80 for water, default radii) were used (short hand notation: B3LYP-D3/def2-TZVP/COSMO).

Reported Gibbs free energies (G) were obtained according to:

$$G = E_{el} + \Delta G_{corr} + \Delta G^{\circ/*}. \tag{2.39}$$

where E_{el} is the electronic energy, ΔG_{corr} the free energy correction obtained from the normal mode analysis, based on the rigid-rotor, harmonic oscillator approximations, using the Turbomole module freeh at a pressure of 1 bar and a temperature of 298.15 K; no scaling factor was applied to the vibrational frequencies. Further, the change in standard state ($\Delta G^{\circ/*}$) from 1 bar to 1 M (resp. 1 bar to 55.6 M for water) is obtained according to:

$$\Delta G^{\circ/*} = RTln(24.4 \cdot c), \tag{2.40}$$

where R is the universal gas constant, T the temperature (298.15 K) and the c the concentration (1 M).

In order to speed-up the simulations the resolution-of-the-identity-density-fitting (RI-J, with its corresponding auxiliary basis sets[229–231]) and the Multipole-Assisted-RI-J[232] (MARIJ) were used.

The choice of those settings is justified by previous studies where it has been shown that the BP86 functional accurately reproduces molecular structures of transition metal complexes.[233–238] In addition, Kwapien *et al.* showed that for cobalt complexes the electronic energies calculated with the B3LYP hybrid functional are comparable with the obtained by coupled cluster singlet doublet (triplet) CCSD(T).[118] In an similar study Kang *et al.* showed that in the case of ruthenium complexes reasonable energetics for the O−O bond formation can be obtained by B3LYP.[239]

2.5.2 CP2K

Some static explicit solvation as well as AIMD employing the BO approximation (BO-MD) have been carried out using the QUICKSTEP method[240] as implemented in the CP2K software package.[241]

KS-DFT was used with a GGA exchange-correlation functional such as BP86,[224,225] PBE[192] or BLYP[198,224] depending on the system and the task at hand. For all simulations the double-ζ valence polarized basis sets (DZVP-MOLOPT-GTH) or its short-ranged version (DZVP-MOLOPT-SR-GTH) together with the corresponding (scalar-relativistic) Goedecker-Teter-Hutter pseudopotentials (GTH)[38,242,243] were used. Further Grimme's D3 dispersion correction was used to account for long-range dispersion interactions.[199]

In some cases, electronic energies were obtained from single point energy calculations of previously optimized structures employing B3LYP.[197,198,244] The simulations were usually performed in the NVT ensemble with a time-step of 0.5 fs, where the temperature was kept constant at 300 K by a Nosé-Hoover chain thermostat.[210,211]

The specific computational settings for each chapter can be found in the corresponding sections.

Chapter 3

Cobalt-based Water Oxidation: a Mechanistic Overview

This chapter is a shorted version of: Schilling, M.; Luber, S.; *Front Chem* **2018**, *6*, 100.[13] If not stated otherwise all figures of this chapter were published in the reference mentioned beforehand.

The aim of this chapter is to give the reader a brief overview of some of the most important contributions in the field of Co-based water oxidation. Due to the scope of this work we focus on contribution focusing on mechanistic aspects of homogeneous WOCs. In order to make the comparison between different catalytic cycles feasible without having to decipher non-standardized reactions schemes, a short-hand notation to describe catalytic cycle in terms of its the elementary reactions is introduced.

3.1 Water Oxidation Mechanisms - Short Hand Notation

We have already introduced the most important aspects of transition metal catalyzed water oxidation in Section 1.2, where a generalized reaction scheme is shown (see Figure 1.1). The catalytic cycle is composed of elementary reactions such as: PT and ET reactions, those might occur consecutively or concerted as PCET. Further, there is the O−O bond formation that might follows one of the many possibilities shown in Figure 1.2. Once O_2 is formed, it has to dissociate (O2DI) from the catalyst in order to allow for the coordination of a new substrate i.e. the association of a water molecule (AQAS), and thereby regenerating the catalytic ground state.

The mechanism of water oxidation catalyzed by a hypothetical $\{M^{II}-OH_2\}^{2+}$ catalyst might be abbreviated by this nomenclature as PCET-PCET-[WNA-PCET]-PCET-[O2DI-AQAS]. Here, the brackets indicate that those elementary reactions were combined in a concerted reaction i.e. that no intermediate state was observed or calculated. Please note, while the brackets represent a concerted reaction in a chemical sense this does not necessarily mean that there is any experimental evidence to support this claim. On the contrary, often there are technical (computational) reasons why certain reactions are assumed to be concerted. Some of those choices are not always intuitive, such as for example: [WNA-PCET] in this case the calculation of a potentially high energy species such as $\{M-OOH_2\}^{3+}$, $\{M-OOH_2\}^{2+}$ or $\{M-OOH\}^{+}$ respectively is avoided. Instead, $\{M-OOH\}^{2+}$ is calculated an intermediate that does not possess a highly acid group and has the same overall charge as the catalytic ground state.

For both real and model systems the exact order of those building blocks is dependent on the nature of the catalyst, i.e. the number transition metal centers and associated oxidation state(s) as well as ligands and environment.

In the following section, the mechanism of several homogeneous Co-based WOCs will be discussed. For other well studied transition metal WOCs we refer to recent reviews.[16,21,22,140,245,246]

3.2 Cobalt-based Water Oxidation Catalysts

3.2.1 Mononuclear WOCs

Corroles

Computational studies on the water oxidation mechanism of Co-based WOCs are rather scarce. Nevertheless, some WOCs were studied with great care, among them are the so-called Hangman-Corroles.[247] Those molecules feature a β-octafluoro corrole (cor) with a linker bearing a carboxyl acid is placed above the face of the macrocycle and can act as a proton acceptor or proton donor; this distinctive structural feature is known as Hangman construct (see Figure 3.1).[248]

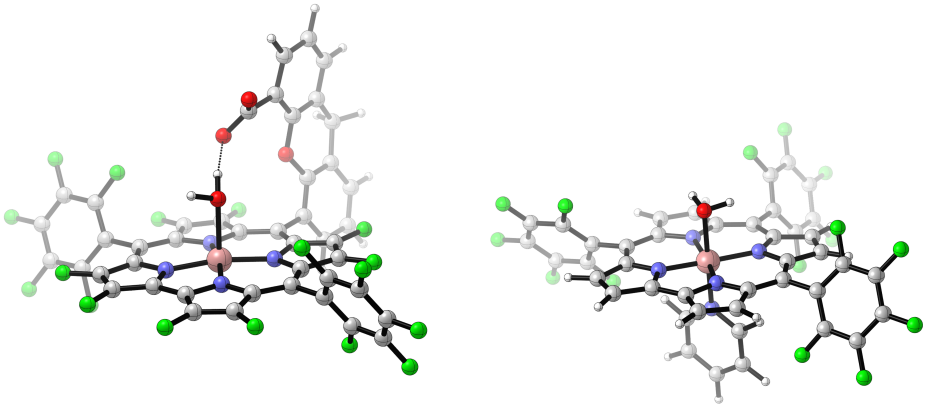


Figure 3.1: Schematic illustration of a Hangman-Corrole (left)[249] and an ordinary corrole complex (right).[250] Reproduced from geometrical data.[249,250]

Based on calculations of reduction potentials, free energies of deprotonations, PCETs, as well as transition states for the O–O bond formation by a WNA and the displacement of molecular oxygen by a solvent molecule, the following reaction mechanism was proposed for the Hangman-Corroles: Starting from a $Co^{III}-OH_2$ state, two consecutive PCETs lead to a reactive state, where a WNA takes place, the carboxylate of the Hangman construct thereby acting as an intra-molecular base and facilitating the O–O bond formation. After an additional PCET, either deprotonation of the carboxylic acid followed by the substitution of molecular oxygen by a solvent molecule or a fourth PCET takes place after which the displacement of the superoxide by a water molecule occurs. Applying the nomenclature established earlier, the mechanism can be described as: PCET-PCET-[(i-PT)-WNA]-PCET-PT-O2DI-AQAS-ET or PCET-PCET-[(i-PT)-WNA]-PCET-PCET-[O2DI-AQAS]. A key factor determining the likelihood for a WNA is the electrophilicity of the oxo/oxyl

ligand, which in turn is governed by the oxidation state of the metal center. Here, the WNA takes place at a formal Co^{V} center. However, calculations indicate radical character on the oxo ligand and on the corrole frame work.[251] Those two non-innocent ligands are formally oxidized instead of the metal center. This results in an electronic structure which cannot be described by standard DFT anymore, therefore the broken-symmetry approach was employed to obtain the energies of the anti-ferromagnetically coupled spins system.

The influence of different transition metals was studied by comparing activation barriers calculated for formal $[M^{V}(cor-COO)(O)]^{-}$ and $[M^{IV}(cor-COO)(O)]^{2-}$ intermediates (M = Co, Mn, Fe, Ru, and Ir).[249] While the $[M^{V}(cor-COO)(O)]^{-}$ state is expected to undergo a WNA,[251] $[M^{IV}(cor-COO)(O)]^{2-}$ represents an intermediate from a different catalytic cycle (lower formal oxidation state before the O–O bond formation) which supposedly starts with the steps PCET-PT-WNA. Since the transition metal center and its oxidation state guide the water oxidation mechanism, a clear preference for the higher oxidation state is expected. It is also important to note that the formal oxidation states not necessarily represent the observed oxidation state, since the ligands are non-innocent.[251] The reported barriers range from 3.6 kcal mol^{-1} for Co^{V} up to 58.3 kcal mol^{-1} for Mn^{IV} for the O–O bond formation.[249] As expected, in general a higher oxidation state favors a WNA attack, which can be rationalized by the stronger electrophilic character of the catalyst. Further they argue that in the series of the 3d transition metals the ability for two electron reduction governs the reactivity. Those two electron reductions are defined as electronic energy difference for the [(i-PT)+WNA] step, which can be thought of as a PT, followed by two single ETs from OH^{-} to the catalyst, and finally an attack of the complex to a OH^{+} [249].

A structural analogue of the Hangman-Corroles without an intramolecular base was investigated by Lei *et al.* (see Figure 3.1).[250] The main structural difference is the assumed octahedral coordination of the metal center compared to a quadratic pyramidal coordination in case of the Hangman-Corroles. Here, two pyridines occupy the coordination sites above and below the corrole plane. The assumed catalytic ground state $[Co^{III}(cor)(OH_2)(py)]$ is reached after the substitution of one pyridine by a water molecule from the solvent. The dissociation of the pyridine thus becomes a gateway step towards the catalytic cycle, which is backed-up by experimental studies. The proposed water oxidation mechanism PCET-PCET-[WNA-PT]-PCET-ET-[O2DI-AQAS] is similar to earlier proposals.[251] The barrier for the O–O bond formation (29.9 kcal mol^{-1}) could be reduced by more than 10 kcal mol^{-1} if an intermolecular base, here OAc^{-} was included into the transition state model. The reactive intermediate is again best described as a $[Co^{III}(cor^{\bullet+})(O^{\bullet+})(py)]$ species rather than a $[Co^{V}(cor)(O)(py)]$. Recently, a series of Co-corroles was reinvestigated employing BS-DFT calculations.[252] Where the presumed $Co^{III}(cor)$ state was found to have significant contributions of an anti-ferromagnetically coupled $Co^{II}(cor^{\bullet})$ state, highlighting the fact that corroles are non-innocent. If those findings are also true for the WOCs discussed earlier, and whether they would have any influence on the reaction mechanism is still under debate.

Polypyridine and Triazole

A common structural feature of most WOCs is the coordination by polydentate nitrogen bearing ligands such as corroles, porphyrines, or pyridine derivatives. An-

other catalyst with such a pyridine ligand scaffold is $[Co^{II}(cPy5)(OH_2)]^{2+}$ (cPy5 = 2,6-(bis(bis-2-pyridyl)-methoxymethane)pyridine) (see Figure 3.2).[253] In contrast to the WOCs described above, it is the first time that the catalytic ground state contains Co^{II} instead of Co^{III}. This has implications on the reaction mechanism: first, the ligand sphere of Co^{II} complexes is prone to ligand exchange, secondly high oxidation states such as Co^{V} can be avoided. On the other side, the lower oxidation state in principle reduces the electrophilicity of the reactive $Co^{III}-O^{\bullet}$ species which in turn makes a WNA mechanism less likely. Nevertheless, the catalytic cycle proposed by Crandell *et al.* follows the standard pattern established before.[254] After two consecutive PCETs, a nucleophilic attack of a hydroxide on the oxyl-species takes place. A subsequent two electron oxidation and deprotonation results in the release of molecular oxygen, and upon coordination of a water molecule the catalytic ground state is recovered (PCET-PCET-WNA(OH)-[PCET+ET]-[O2DI-AQAS]).

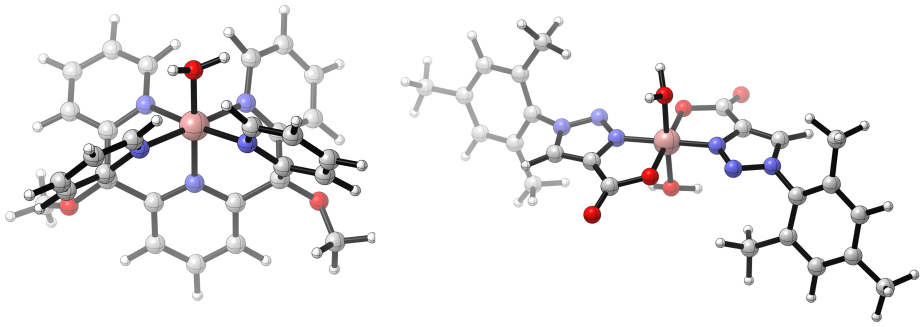


Figure 3.2: Schematic illustration of $[Co^{II}(Py5)(OH_2)]^{2+}$ (left),[254] and $[Co^{II}(TCA)_2(H_2O)_2]$ (right).[255] Reproduced from geometrical data.[254,255]

Further the nature i.e whether the WNA is an intra- or intermolecular process was investigated.[254] This is crucial, since in principle a hydroxide molecule could coordinate to the metal center prior to the nucleophilic attack onto the oxyl-species, a possibility which has often been overlooked in similar studies. Because the metal center is already coordinatively saturated at the reactive intermediate ($Co^{III}-O^{\bullet}$), and ligand exchange reactions are expected to be unlikely at highly oxidized cobalt centers, decoordination of a pyridine at an earlier stage of the catalytic cycle is required. The barrier for the substitution of an equatorial pyridine by a hydroxide in the catalytic ground state was found to be 18.7 kcal mol^{-1}, which might be overcome under experimental conditions. The barrier for the WNA itself was found to be smaller for an intramolecular (21.1 kcal mol^{-1}) than for an intermolecular mechanism (29.9 kcal mol^{-1}), but if the relative stability of the oxyl-species is taken into account (8.8 kcal mol^{-1} in favor of the η^5-Py5 species) both mechanisms become energetically equivalent. In this case, the intermolecular mechanism is expected to dominate because of the energy penalty associated with the coordination of hydroxide (structures where a pyridine ligand is dissociated are higher in free energy - see Figure 3.3).

Recently, a mononuclear Co-catalyst with a completely different ligand frame

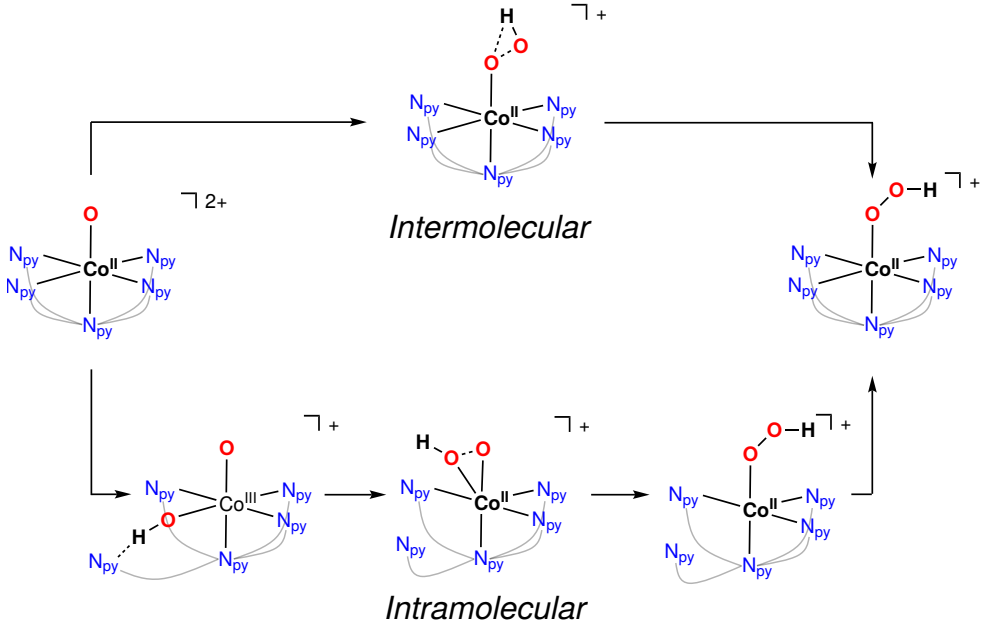


Figure 3.3: Reaction mechanism of the O–O bond formation catalyzed by $[Co^{II}(Py5)(OH_2)]^{2+}$. Note the pentadentate ligand (blue) is simplified for illustrative purposes only (see Figure 3.2). Reproduced from data published by Crandell *et al.*[254] (*unpublished figure*).

work has been reported (see Figure 3.2).[255] This is composed of bidentate triazolcarboxylates that coordinate with one oxygen of the carboxylate as well as one nitrogen atom of the triazol to the metal center. The final complex is composed of two of those ligands in a *trans* configuration $[Co^{II}(TCA)_2(H_2O)_2]$ (TCA = 1-mesityl-1,2,3-1H-triazole-4-carboxylate). Based on experimental data the following reaction mechanism was proposed: PCET-PCET-[WNA-PCET]-PCET-[O2DI-AQAS]. The two alternative pathways, i.e. RC and a bifunctional mechanism, were also considered. However, a RC mechanism could be excluded by experiments, and the bifunctional mechanism was ruled out since there is no suitable hydrogen acceptor in close proximity to facilitate a concerted mechanism.

3.2.2 Dinuclear WOCs

Polypyridine

Moving on from mononuclear Co-based WOCs to dinuclear catalysts, only a few examples of mechanistic studies can be found in the literature. An example is $[Co^{III}{}_2(TPA)_2(OH)_2]^{4+}$ (TPA = tris(2-pyridylmethyl)amine) (see Figure 3.4).[256] The main structural feature of those complexes are the two bridging hydroxyl ligands that are not only crucial for the formation of the dimeric structure, they are also the substrate for the O–O bond formation. The proposed mechanism can be abbreviated as PT-[ET-PCET]-(i-RC(μ-O))-[AQAS-PT]-ET-ET-O2DI-AQAS-PT-PT, where the first AQAS-PT refers to an insertion of a H_2O molecule in between the metal centers forming a new hydroxyl bridge. Employing DFT calculations the spin state of the catalytic intermediate that supposedly undergoes the O–O bond formation $[Co^{III}{}_2(TPA)_2(O^{\bullet})_2]^{4+}$, was characterized.[256] For this species, the triplet spin state

was found to be the most stable one, suggesting the formation of two μ-oxyl species. Formation of the O−O bond leads to a formal superoxide which is coordinated in a side-on (η^2) fashion to each of the two metal centers. Upon coordination of a solvent molecule the coordination mode changes to endo-on (η^1), and subsequent oxidation leads to the release of molecular oxygen.

Another dinuclear catalyst - $[(Co^{III}(trpy))_2(\mu\text{-bpp})(\mu\text{-OO})]^{3+}$ (trpy = 2,2′;6′:2″-terpyridine; bpp^- = bis(2-pyridyl)-3,5-pyrazolate) was found not only to be able to oxidize water, but also to perform the reverse reaction as well as to catalyze the water reduction reaction.[257–259] From an experimental point of view, this aspect is extensively discussed in literature.[260,261] Unlike most dinuclear WOCs, the catalyst does not possess any bridging oxo-ligands, the pyrazol subunit of the bbp ligand takes over this role. The other coordination sites of the octahedrally coordinated metal centers are occupied by trpy ligands as well as solvent molecules which are the substrate for the water oxidation (see Figure 3.4). The proposed mechanism for water oxidation is initiated by a formal $[(Co^{III}(trpy)(OH_2))(\mu\text{-bpp})(Co^{III}(trpy)(OH))]^{4+}$ species, the latter undergoes two PCETs, a deprotonation, an O−O bond formation by RC, followed by regeneration of the initial state by H_2O insertions and further oxidations. In summary: PCET-PCET-[PT-(i-RC)]-ET-AQAS-PCET-[O2DI-AQAS] (at pH = 2). A remarkable feature of the proposed catalytic cycle is the fact that even though the resting state is composed of two Co^{III} centers, none of them is expected to be oxidized to a formal oxidation state higher than Co^{IV} (formal Co^{III}-oxyl). The mechanism is backed up by exhaustive calculations of dehydrogenation free energies (PCETs), redox-potentials, and p$K_a$values for different conditions (pH values), as well as activation energies for the O−O formation, the change of the coordination mode of O_2 from η^2 to η^1, and the O_2 release.

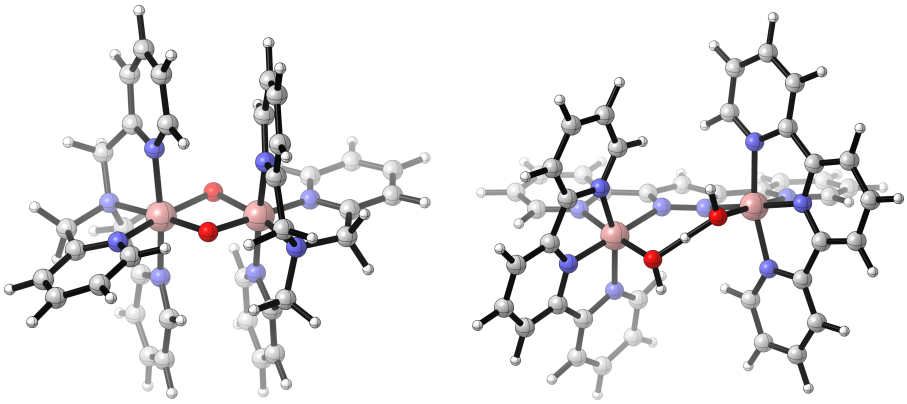


Figure 3.4: Schematic illustration of $[Co^{III}{}_2(TPA)_2(\mu\text{-}O)_2]^{2+}$ (left),[256] and $[Co^{III}(trpy)\}_2(\mu\text{-bpp})(OH_2)(OH)]^{4+}$ (right).[259] Reproduced from geometrical data.[256,259]

Model System

Possible water oxidation mechanisms for a dinuclear model system $[(Co^{III}(OH_2)_2(OH)_2)_2(\mu\text{-}OH)_2]$ were studied by Fernando and Aikens.[120] Obviously there are many protonation isomers, therefore the complexity quickly

increases along the possible reaction path. For example, there exist 14 possibilities for the first PCET, then for each of the isomers obtained after the first PCET, there are 13 possibilities for the second PCET, and so on. The stability of those isomers is largely determined by their intramolecular hydrogen bonding network which may be too favored due to the use of a solvent continuum model (see Figure 3.5). With the given structures there are plenty of possible pathways for the O–O formation. The ones described in the study can be grouped into intra- and intermolecular WNAs, where the nucleophile is either a water molecule, a μ-OH or a geminal OH. The WNA pathway (PCET-PCET-PCET-WNA-PCET-O2DI-AQAS) is reported to be thermodynamically favorable over the two possible RC mechanisms, but no barriers were calculated which support this argument.[120]

The active species for all three possible nucleophiles is reported to be $Co^{V}-O^{\bullet}$ (formal $Co^{VI}=O$) obtained after three consecutive PCETs, all of which have to happen at the same metal center. While the authors do not attempt to assign formal oxidation states to the individual metal centers, the nomenclature of the oxyl as well as the reported spin multiplicity clearly suggest a single-site mechanism. In this light, the question arises why the second metal center is even necessary, a question which was not discussed in the study. So why is the dinuclear model system still important? It might be envisioned to be the smallest building block of Co_xO_3, which is known to be a potent heterogeneous WOC. Cobalt oxide is also expected to be the main component of the famous CoP_i catalyst reported by Kanan and Nocera.[262] The next larger subunit – trinuclear complexes complexes such as $[Co_3(O)(OH)_2(OAc)_3(py)_5]^{2+}$ – was found to decompose under catalytic conditions and therefore were not studied in great depth.[119,263]

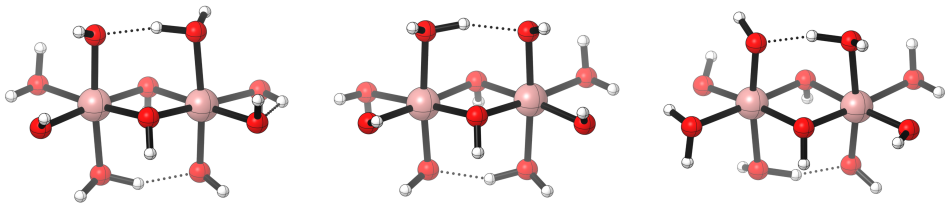


Figure 3.5: Schematic illustration of three $[(Co^{III}(OH_2)_2(OH)_2)_2(\mu\text{-}OH)_2]$ isomers. Electronic energy difference given relative to the left structure, 0.0 eV (left),[120] 0.12 eV (middle),[120] 0.21 eV (right).[120] Reproduced from geometrical data.[120]

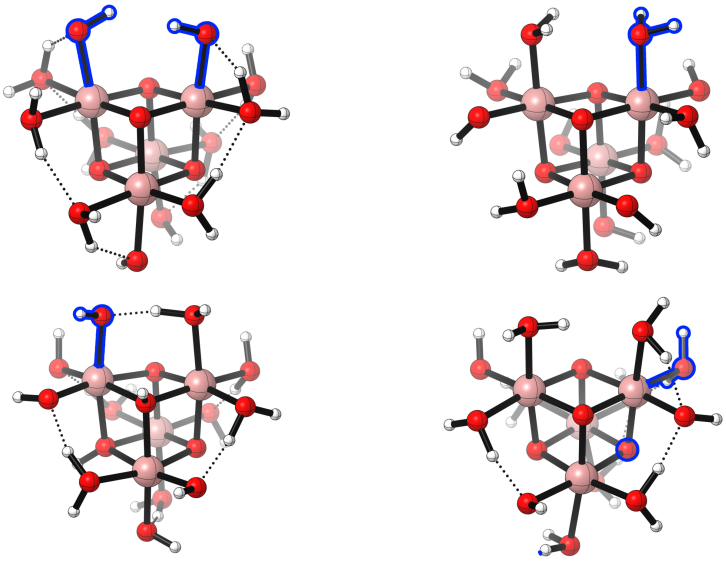


Figure 3.6: Schematic illustration of the [($Co^{III}(OH_2)_2(OH)(\mu^3$-O))$_4$] model systems (top, left),[117] (top, right),[118] (bottom, left),[119] and (bottom, right).[120] If indicated in the original paper, hydrogen bonds are shown. Further, the 'active' ligands, i.e. the ones which undergo the O−O bond formation, are highlighted in blue. Reproduced from geometrical data.[117–120]

3.2.3 Polynuclear WOCs

Model System

Catalytically active and stable polynuclear WOCs with four centers exhibit a cuboidal structure (see Figure 3.6). Like for the dinuclear species, there is a large number of possible protonation isomers.[120] The thermodynamically most stable ones were selected for further investigations. Again different pathways for the O−O bond formation - a WNA, an i-WNA, i-RC(geminal), and i-RC were investigated. Here it is worth to mention that the bridging oxygens are not protonated (μ-O) in their model (see Figure 3.6). From a thermodynamical point of view, radical coupling between an oxyl and a bridging oxygen turned out to be favorable (PCET-PCET-(i-RC(μ-O))-AQAS-[AQAS-PCET]-O2DI-PCET). A different protonation isomer, where the two hydroxyl ligands of Co are on the same site was found to be energetically unfavorable. Nevertheless, it might be a more realistic model for Co_xO_3 surface. For those models, an additional pathway arises, namely i-RC between two oxo species. All discussed pathways share a common starting sequence of reactions – two consecutive PCETs, leading to $Co^{IV}-O^{\bullet}$ or $Co^{V}=O$. The authors use a slightly misleading nomenclature ($Co^{V}-O^{\bullet}$ radical), which would imply a formal oxidation state of Co^{VI} for the Co-oxo species. However such a high oxidation state can only be reached after three consecutive oxidation steps (see dimer). As the electrophilicity of $Co^{V}-O^{\bullet}$ and $Co^{IV}-O^{\bullet}$ is expected to be significantly different, it might serves as a rational for the preference of the RC mechanism over a WNA. Even though this model system oversimplifies aspects of a real system such as solute-solvent interactions, it is one of the few studies which shows the complexity arising from different protonation sites. For real catalysts, those considerations

might not be as important as for the model system since the number of acidic / basic sites is often limited. Nevertheless, there might be the possibility of ligand decoordination or substitution which might result in a high number of protonation isomers. In order to overcome the flaws of implicit solvation models, in particular the treatment of hydrogen bonding between the solute and the solvent, the same model system as has been used by Fernando and Aikens,[120] was placed in a box of 216 water molecules and treated on a QM/MM level of theory.[117] A protonation state where two hydroxyl ligands are in a *syn* configuration was chosen as the reactive intermediate (see Figure 3.6). For those systems the activation barriers were calculated in order to discriminate certain pathways, among them was a WNA, an i-WNA(OH), as well as i-RC between two bridging oxo-ligands. Energetically all of those were found to be inferior to the standard i-RC between two $Co^{III}-O^{\bullet}$. The reaction path might be summarized as PCET-PCET-(i-RC)-[AQAS-PCET-PCET]-[O2DI-AQAS]. Like in earlier reports only two PCETs were found to be necessary in order to arrive at catalytically active oxyl species ($Co^{III}-O^{\bullet}$).[120] It is important to note that the spin state of the intermediate prior to the O−O bond formation is an open-shell singlet (i.e. anti ferromagnetic-coupling between the metal centers). However, the ferromagnetic triplet state was found to be only 1.6 kcal mol^{-1} higher in electronic energy. An interesting observation is also that the activation energy for an i-WNA by a hydroxyl species was found to be much higher than the one for an i-RC mechanism. This might raise the question how large the barriers for an i-RC are in case of the dinuclear model complex discussed by Fernando and Aikens.[120] In their conclusion Wang and Van Voorhis challenge one of the key assumptions most experimentalists and theoreticians make,[117] by proposing that either the addition of a water molecule or an intramolecular PT might be rate-limiting rather than the O−O formation. While based on experimental conditions (e.g. the presence of strong oxidants or bases) it often can be justified that ET and intermolecular PT steps are irreversible and not rate-limiting, this is not necessarily true for chemical steps (e.g. O−O bond formation, O_2 release, or intramolecular PTs).

In a similar study, the same model system [$(Co^{III}(OH_2)_2(OH)(\mu^3\text{-}O))_4$] (see Figure 3.6) and the smallest possible subunit of it,an aqua ion [$Co^{III}(OH_2)_3(OH)_3$], were used in order to answer the question whether the choice of the exchange-correlation density functional has an effect on the structure of the intermediates and their reaction energies.[118] For both systems classical single site mechanism (PCET-PCET-WNA-PCET-PCET-O2DI-AQAS) was assumed for both model systems. When comparing the hybrid functionals B3LYP and PBE0 with the GGA functional PBE, significant deviations were found in terms of the energetic ordering of different isomers, i. e. the thermodynamically favored product when the $\{H_2O-Co-OOH\}$ fragment undergoes a PCET (superoxo $\{H_2O-Co-OO\}$ or hydroperoxo fragment $\{HO-Co-OOH\}$). To further validate the density functionals, the electronic energy differences for the mononuclear Co-complex were compared with ones obtained from CCSD(T) calculations. A good agreement of both hybrid functionals, from a quantitative and qualitative point of view, with the CCSD(T) results was found. Such benchmark studies are important, in particular since the tetranuclear systems of interest are usually too complex for the application of a more accurate level of theory beyond DFT.

A different protonation state of the same model system as discussed above, where two bridging oxo ligands bear a proton in the resting state, was investigated by

Li and Siegbahn (see Figure 3.6).[119] This protomer was found to be considerably lower in free energy than the models discussed above (see Figure 3.6).[117,120] The minimum energy reaction pathway proposed is again similar to the ones already discussed. Starting from the all Co^{III} resting state, two consecutive PCETs take place. Interestingly, the successive oxidation of a single cobalt center is favored over the oxidation of two equivalent centers. The resulting Co^{V}=O then undergoes a WNA, followed by two PCETs and the release of molecular oxygen (PCET-PCET-[WNA-PCET]-PCET-[O2DI-AQAS]).[119] Other reaction pathways were found to be energetically unfavorable because they contain either high energy intermediates or the barrier for the O−O formation was found to be too large. Even larger building blocks mimicking the heterogeneous Co_xO_3 catalyst were furthermore investigated using AIMD and DFT+U.[264] Due to the focus on homogeneous catalysis we refrain from an in-depth discussion them.

Co^{III} Cubanes

Besides the model systems, there are also actual catalysts featuring a cuboidal core. Among them are the Dismukes-cubanes $\{Co^{III}{}_4O_4\}$ ($[Co^{III}{}_4(\mu^3\text{-O})_4(\mu^2\text{-OAc})_4(py)_4]$, OAc = acetate) that were already discussed in Section 1.3.1. Here, only a brief description of the mechanistic proposals rather than a full recapitulation of all the related studies is given.

Due to the structural similarity of their model system and the Dismukes-cubane, Li and Siegbahn used the same protocol to describe the first steps of the mechanism.[119] The sequence of which turned out to be identical for both the model system and the real catalyst. The close proximity of an OAc^- to the active site Co^{V}=O, in case of the Dismukes-cubane, allowed it to act as an intramolecular base, facilitating the WNA, resulting in a lower overall barrier (see Figure 1.4).

Usually the isolation and characterization of catalytic intermediates is very challenging. Despite all odds, in the case of the Dismukes-cubanes the isolation of the once oxidized intermediate $\{Co^{III}{}_3Co^{IV}O_4\}^+$ was successful. This observation inspired Nguyen *et al.* to a mechanistic proposal where $\{Co^{III}{}_3Co^{IV}O_4\}^+$ disproportionates in order to form a highly reactive Co^{V}=O species.[123] Except for the formation of said intermediate the proposed pathway (ET-PT-ET-WNA(OH)-PCET-[ET-A2AS(OH)-O2DI]) is similar to the one suggested by Li and Siegbahn for their model system (PCET-PCET-[WNA-PCET]-PCET-[O2DI-AQAS]),[119] as well as the one proposed by Kwapien *et al.* (PCET-PCET-WNA-PCET-PCET-O2DI-AQAS).[118]

As mentioned in Section 1.3.1, the speciation of $\{Co^{III}{}_4O_4\}$ in solution remains unknown, as a consequence various structural proposals exist (see Figure 1.4). In particular the association of second hydroxyl ligand is discussed exhaustively.[122,124] In their latest study Nguyen *et al.* proposed based on both experimental evidence as well as calculations a *syn* configuration of the two hydroxyl ligands. Adopting their early mechanistic postulate to the newly identified intermediates, they ended up with ET-PT-ET-(i-WNA(OH))-PCET-[AQAS(2*OH^-)-O2DI-ET], where the hydroxyl coordinated to a Co^{III} center nucleophilically attacks Co^{V}−oxo.

Recently, the single-site mechanisms discussed beforehand have been challenged by an experimental study where they were able to characterize the doubly oxidized species $[Co^{III}{}_2Co^{IV}{}_2O_4(\mu^2\text{-OAc})_4(Py)]^{2+}$.[126] Based on cyclic voltammetry measurements, X-ray absorption spectroscopy as well as BS-DFT calculations, they con-

cluded that the species contains an anti-ferromagnetically coupled $\{Co^{III}{}_2Co^{IV}{}_2\}$ core rather than a $\{Co^{III}{}_3Co^{V}\}$ core. If those findings also apply to intermediates where some of the acetate ligands have been replaced by aqua or hydroxyl ligands, then a multi-site mechanism such as the i-RC proposed by Wang and Van Voorhis should be favorable over a single-site mechanism.[117]

Co^{II} Cubanes

Besides the Co^{III}-based Dismukes-cubanes there is the family of Co^{II} cubanes that has been reported by Patzke *et al.*[38,115,116] Those WOCs have already been discussed exhaustively in Section 1.3.2. Here we only summarize the proposed mechanisms in terms of the established nomenclature. For $\{Co_4(pyMeO)_4\}$, despite the fact that substitution of an acetate ligand by a solvent molecule was found to be thermodynamically unfavorable,[115] both a WNA and a i-RC mechanism, were considered.[131] The two mechanisms might be summarized as PCET-PCET-[WNA-PCET]-PCET-[O2DI-AQAS] and PCET-PCET-PCET-PCET-(i-RC)-[O2DI-AQAS]. Both the thermodynamics of the whole catalytic cycle as well as the activation barrier were found to be favorable in case of the i-RC mechanism. However, the direct comparison of the two pathways is flawed due to the fact that they assume a different reactive intermediate and compare two fundamentally different activation barriers. In case of the i-RC mechanism O_2 is formed, while the products of a WNA are OOH and H_3O^+. Nonetheless, both proposed pathways have in common that the highest formal oxidation state is Co^{IV} rather than Co^{V}, which has been suggested for several mononuclear WOCs.[253,255]

For the second generation of Co^{II}-base cubanes $\{Co_3Ln(pyMeO)_4\}$ an even more exhaustive study has been conducted.[38,67] The mechanistic proposal assumes a single-site reactivity analogously to $\{Co_4(pyMeO)_4\}$. Due to the slightly different ligand environment where the reactive cobalt center bears both a water and a hydroxyl ligand two variations of a WNA have been proposed (pathway A and B, see Figure 1.6) - both might be summarized as PCET-PCET-[WNA-PCET]-PCET-[O2DI-AQAS]. While from a thermodynamic point of view pathway B turned out to be favorable, pathway A was found to be kinetically more feasible. This might be attributed to the fact that formation of hydronium ion is energetically unfavorable (see Section 1.3.2). To some extent this observation might also be related to the fact that NEB calculations are not optimally suited to describe process involving solvent molecules.

The flexibility of the cubane core found during those studies led to the discovery of the 'open'-cubane configurations. The thermodynamics of the water oxidation reaction catalyzed by the Co^{II} cubane was computed assuming the same PCET-PCET-[WNA-PCET]-PCET-[O2DI-AQAS] pathway. No improvement of the overall improvement was found when comparing the thermodynamics of the 'open' and 'closed' configurations.[133] For more details see Chapter 4.

Polyoxometalates

Besides the organometallic cubane clusters, there are also carbon-free Co^{II} WOCs, such as the polyoxometalates (POMs).[265] Those complexes are composed of a tetra-cobalt oxide core which is sandwiched by two lacunary polytungstate cages (XW_9O_{34} (X = P or V)) forming $[Co_4(H_2O)_2(XW_9O_{34})_2]^{10-}$ (see Figure 3.7). The cobalt atoms do not form a cuboidal core but rather a mimic of a cobalt oxide layer.

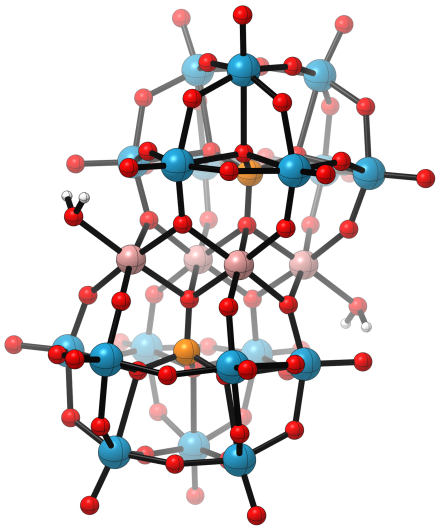


Figure 3.7: Schematic illustration of $[Co_4(H_2O)_2(XW_9O_{34})_2]^{10-}$ (X = P or V) POM.[265] Extension of the color code: *orange* – phosphorous or vanadium, *pale blue* - tungsten. Reproduced from geometrical data.[265]

As a consequence, only the two terminal metal centers possess aqua ligands, which limits O–O bond formation to inter – rather than intramolecular pathways. Since a bimolecular pathways seems unlikely due to the sheer size of the POM, only one option is left: PT-ET-PCET-[WNA-(i-PT)-ET]-PCET-[O2DI-AQAS] and concerted variations of it. Free energy calculations favored the sequence stated before which is identical to the mechanism proposed for the Co^{II}-cubanes. The only difference is the capability of bridging oxo moieties (Co–O–W) to act as a proton acceptor, and thereby facilitate a WNA.

3.3 Summary and Conclusion

In the following some general conclusions on the calculated mechanisms are presented. We want to emphasize that the goal of this summary is not to judge whether the proposed mechanisms are in agreement with experiments or not. The purpose is mere to compare the mechanisms from a theoretical point of view. Comparing the proposed mechanisms for a certain type of ligand framework and nuclearity, we find some correlations that might help to elucidate the reaction mechanism of yet undiscovered catalysts. However, there might also be the chance that up to now the proposed mechanisms are somewhat biased by the previous studies and in particular by the computational protocols applied.

There is general agreement on the first steps of the mechanism of mononuclear cobalt WOCs, which are supposed to undergo two PCETs before the O–O bond is formed by a WNA. Depending on the WOCs initial oxidation state, two electron transfers lead to either a formal Co^{IV} or Co^{V} species. All potential Co^{V} complexes are bearing non-innocent ligands that potentially are oxidized instead of the metal center, hence a formal oxidation state of Co^{IV} might be a sufficient requirement for water oxidation. Further, for the discussed mononuclear catalysts a bimolecular RC

mechanism (I2M) was either excluded by experiments or rendered unlikely due to the strict octahedral coordination mode of the cobalt center as well as the sterically demanding ligands. An extension of the coordination sphere as observed for certain ruthenium-based catalysts seems unlikely for the used ligand-frameworks, which are rather confined in their flexibility.

The picture is similar for dinuclear catalysts in terms of the catalytically active species and formal its oxidation state (Co^{IV}=O). However, when it comes to the O−O bond formation, i-RC becomes a respectable alternative to a WNA, since bulky ligands are hardly the limiting factor for intramolecular reactions. Regarding the Co^{III} cubodial model systems there is consent among all authors: two PCETs are followed by the O−O bond formation. The preferred mechanism for the bond formation of course depends on the protonation state (i.e. protonated μ-O, oxo-ligands on the same face of the cube etc.) which in turn is dependent on the solvation model. The proposed mechanism for the *Dismukes-Cubane* changed slightly over the past years: from a WNA, over a WNA(OH), to an i-WNA(OH). The reason for the changes are mainly routed in the uncertainty about the exact structure (i.e. substitution of OAc-ligands by H_2O or OH^-) as well as the protonation state of the species in solution. Nevertheless, in general the mechanism is still consistent with the ones proposed for model systems. This is also true for the Co^{II}-cubanes, which both supposedly follow a reaction scheme analogous to the one proposed for Co^{III} model systems. However, also here the authors have explored a variety of structures depending on which and how many of the OAc-ligands were exchanged. The non-cubodial cobalt-core of the Co^{II}-POM was found to follow a similar water oxidation mechanism as the Co^{II}-cubanes, since all those catalysts have in common that the number of metal centers accessible by solvent is limited.

From all those studies we can draw guidelines how to approach mechanistic studies of novel catalysts and how to improve current computational approaches. Of uttermost importance is the choice of an appropriate model system. Since the reaction takes place in the condensed phase, special attention has to be given to possible ligand substitutions and protonation state(s), the latter are in particular dependent on the experimental conditions, where the pH usually is controlled by buffers. But not only the catalyst and buffer ions are part of the reaction solution, there are also counter ions (e.g of the catalyst or the oxidant) and chemical oxidants as, in the case of photocatalytic water oxidation, photosensitzers and sacrificial electron acceptors. If and how those other molecules influence the reaction mechanism is often not directly obvious and might be worth to be investigated.

Chapter 4

Discovery of open cubane-core Structures for bio-mimetic $\{Co_3Ln(pyMeO)_4\}$ Water Oxidation Catalysts

This chapter is a shorted version of: Schilling, M.; Hodel, F. H.; and Luber, S.; *ChemSusChem* **2017** *10*, 4561–4569.[133] If not stated otherwise all figures of this chapter were published in the reference mentioned beforehand.

The first biomimetic WOC based on a Co^{II}-cubane motif, $[Co_4(hmp)_4(\mu\text{-}OAc)_2(\mu^2\text{-}OAc)_2(H_2O)_2]$ (hmp = 2-(hydroxymethyl)-pyridine) $\{Co_4(pyMeO)_4\}$, was reported by Evangelisti *et al.*[115] In a subsequent study, they were able improve the catalytic activity of the WOC by introducing a redox inert Ln^{3+} (Ln = lanthanide) cation mimicking Ca^{2+} found in nature's OEC.[38,115] Hodel and Luber studied those two catalysts in depth in order to elucidate the reaction mechanism governing water oxidation (overview of their results is given in the Sections 1.3.2). In the course of those studies they identified structures with a distorted cubane core.[67]

In this chapter, we set out to further explore this phenomenon. Starting from the formerly reported 'distorted' cubane core structures (see Figure 4.1),[67] additional local minima on the potential energy surface were located where the structures feature Co1−O1 and Co1−O3 bonds which are even more elongated (see Figure 4.1). In the following, we discuss the thermodynamics of water oxidation catalyzed by these 'open' cubane core structures. A brief summary of the key findings of this chapter is given in Section 1.3.3.

4.1 Computational Methodology

We employed the same methods and codes as previously applied to $\{Co_3Er(pyMeO)_4\}$, $\{Co_3Tm(pyOMe)_4\}$ and $\{Co_4(pyMeO)_4\}$.[38,67,131] Those are, BP86-D3/DZVP-MOLOPT-GTH for energy and force calculations and B3LYP-D3/DZVP-MOLOPT-GTH for electronic energies within the program package CP2K[241] (see 2.5.2). As we give here only a brief justification of the computational settings here, we ask the reader to review the literature mentioned beforehand.

The initial atomic coordinates of the S0 state were obtained from a snapshot of a DFT-based MD trajectory where all water molecules except for the 68 closest to the catalyst were deleted. The structures for the S1-S4 states were subsequently gener-

ated from their respective previous state to ensure that the solvent molecules were in comparable positions.[67,131] For this work only one of the two possible configurational isomers of the S2 state is of importance (isomer b) since no open cubane core structures were found for the other isomer.[67] We thus will not distinguish between them in the remaining part of the manuscript. Single point energy calculations were performed on both the geometry optimized system including the catalyst surrounded by 68 water molecules and the catalyst with the same structure but without explicit solvent molecules. All calculations using CP2K were performed in a 30^3 $\AA^3$ simulation box.

The energy cutoff for the auxiliary plane wave expansion of the charge density of was set to 500 Ry. Due to the lack of experimental data, ferromagnetic coupling between the Co and Ln centers was assumed for all catalytic states. The proposed mechanism only involves one active Co center, and alterations of the electronic configurations were exclusively performed at this center. The geometries of all intermediates were optimized for all reasonable spin multiplicities. Then, the electronic structures were analyzed in terms of Mulliken spin populations and frontier (canonical) orbitals. Mulliken spin populations depend on the basis set and in general do not converge to the basis set limit,[266] still they have been routinely used.[118,267]

The differences in free energy between the catalytic intermediates and the ground state ΔG_{i-0} were calculated following a protocol by Nørskov *et al.*[268] which has already been successfully applied to the Co^{II}-based cubanes and other catalysts.[67,70,131,269,270] This approach is based on the assumption that all oxidation processes are PCETs and that these reactions are essentially barrier free. Hence, contributions of the proton and the electron are not treated independently. ΔG_{i-0} is calculated as

$$\begin{aligned}\Delta G_{Si-S0} = &\left(E_{Si} + \frac{i}{2}E_{H_2} - E_{S0}\right) + \left(E_{ZPE,Si} + \frac{i}{2}E_{ZPE,H_2} - E_{ZPE,S0}\right) \\ &+ i\left(\Delta H - T\Delta S\right), i = 1, ...4\end{aligned} \tag{4.1}$$

E_i is the electronic energy of state i, and $E_{ZPE,i}$ is the zero-point-energy. The thermal corrections for the different intermediates of the catalysts were assumed to cancel out. Therefore, only correction terms for molecular hydrogen containing $\Delta S = 0.016\,\mathrm{kcal\,K^{-1}\,mol^{-1}}$ as well as $\delta H = 1.153\,\mathrm{kcal\,mol^{-1}}$, i.e. half the entropy and enthalpy of molecular hydrogen at standard conditions, are included.[271]

The computations of the minimum electronic energy paths and barriers with explicit solvation and the CP2K package were carried out with the NEB method[44] (improved tangent method,[43] climbing image NEB every 5 steps), each pathway consisting of 8 frames. All set-up parameters for the NEB simulations have been selected to be identical to the ones chosen for the geometry optimizations. For all frames along the optimized reaction pathway electronic energies were obtained employing the B3LYP functional. The zero-point energies were obtained from normal mode analysis in the harmonic approximation. Due to the high computational effort, such calculations were performed in TURBOMOLE 7.01[206] with the COSMO[204] instead of explicit solvation (see Section 2.5.1 for a detailed description of those settings). In order to investigate the effects of the explicit solvation shell, all structures optimized in CP2K (with explicit solvation) were re-optimized in TURBOMOLE using COSMO and no explicit water molecules. It is important to note that the novel conformers

would not have been detected without the appropriate initial geometries obtained from CP2K including the explicit first solvation shell.

The COSMO energies were further refined using the DCOSMO-RS procedure as implemented in TURBOMOLE based on the provided σ-potential files for water at 25°C.[64]. Geometry optimizations carried out for a few selected structures using DCOSMO-RS were virtually identical with the structures obtained using COSMO. However, recently DCOSMO-RS was reported to significantly improve activation and reaction energies of organic reactions in protic solvent.[272] The two-center shared electron numbers (SENs) were calculated via population analysis based on occupation numbers as implemented in TURBOMOLE.[273] The kinetic barrier for the cubane core opening was evaluated by calculating an approximate reaction path for this process employing the WOELFLING script with standard settings as defined by TURBOMOLE.[42,274–276] Molecular structures were visualized using VMD or CYLview.[277,278]

4.2 Results and Discussion

4.2.1 Open Cubane-Core Structures

In our previous work we reported an energetically possible 'distorted' conformation for the S2 state of $\{Co_3Ln(pyMeO)_4\}$(see Figure 4.1).[67] Those structures sparked our interest since they might be part of a different catalytic cycle than the one presented by Hodel and Luber.[67] We therefore investigated this phenomenon further. Starting from the previously described structures, we found even more distorted structures – for the sake of brevity in the following referred to as 'open' structures in contrast to 'closed' ones featuring a closed cubane core - which can be used to build a complete catalytic cycle in a consecutive manner. The open structures represent minima on the potential energy surface and exhibit elongated Co1–O1 and Co1–O3 bonds that cause Co1 to change its coordination geometry from octahedral to tetrahedral (see Figure 4.1). The larger Co1–O distances allow the water ligand of Co1 to form an intramolecular hydrogen bond with O1, that further stabilizes the structure. The open cage motif is reminiscent of observations made for the OEC, which is also known to form open and closed cubane core structures.[101,132,279–281] Moreover, the reduction of the CN is similar to a recent postulation of a five-fold coordinated Mn^{IV} intermediate in the water oxidation reaction catalyzed by the OEC.[282] However, it is important to note that the open/closed conformers of the OEC and $\{Co_3Ln(pyMeO)_4\}$ are significantly different. In the case of the OEC, an oxygen atom is pulled-out of the cubane cage by the dangling manganese opening up the cubane core. The catalysts at hand, $\{Co_3Ln(pyMeO)_4\}$, do not possess a fourth redox-active transition metal center, further the oxygen atoms which form the cubane core are a fundamental part of the hmp ligand and therefore restricted in their spatial movement. The covalent linkage of those oxygen atoms to carbon backbone of the ligands is the main reason why we do not expect their involvement in the oxygen-oxygen bond formation process, as considered for other model catalysts.[120] In contrast to the OEC where the bridge-site oxygen's can be protonated,[283] no such involvement can be found for $\{Co_3Ln(pyMeO)_4\}$ because they are coordinately saturated. Additionally a geminal oxo-oxyl coupling pathway can be excluded since a cis-dioxocobalt center would be a prerequisite for such a pathway.[122,284,285] Existence of such a species is questionable due to the requirement of a high oxidation state.

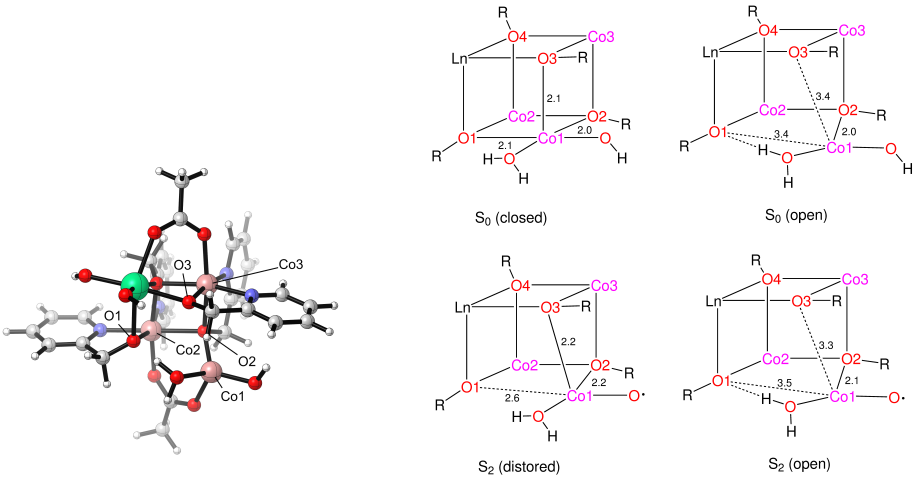


Figure 4.1: Open cubane core structure of {$Co_3Ln(pyMeO)_4$} in the S0 state used in this study. The cobalt centers as well as some key oxygen atoms are labeled. Note that the ligand framework is identical to the one used in our previous study,[67] but not to the reported crystal structure due to ligand exchange (left).[38] Schematic representation of the open and closed cubane core structures of {$Co_3Ln(pyMeO)_4$} in the catalytic ground state S0 (top, right). In particular, the Co1–O1 and Co1–O3 bonds are significantly elongated in the former. The same is true when comparing the distorted and open conformers of intermediate S2 (bottom, right). Distances are given in Å.

4.2.2 Catalytic Water Oxidation

Possible water oxidation pathways based on a WNA mechanism are given in Figure 4.2. In the inner circle, water oxidation catalyzed by the known closed structures is visualized. The outer circle represents the open forms newly presented in this manuscript. We find the open structure of the ground state (S0) to be thermodynamically more stable than the closed one, as indicated by the difference in zero-point corrected electronic energy. This trend is reversed for the first intermediate (S1) where the closed structure is more stable. However, the energy difference is less than 5 kcal mol^{-1}, which indicates - given a sufficiently low energetic barrier in between – that the open and closed structures potentially coexist. For the S0 and S1 states, both catalysts {$Co_3Er(pyMeO)_4$} and {$Co_3Tm(pyOMe)_4$} show a comparable behavior. An exception to this trend is the second intermediate (S2) where the open structure is thermodynamically more stable in case of {$Co_3Er(pyMeO)_4$} while {$Co_3Tm(pyOMe)_4$} clearly favors the closed structure. The difference can be rationalized by comparing some selected bond lengths, namely Co1–O1, Co1–O2, and Co1–O3. Those are within 0.1 Å of one another, when comparing structures of {$Co_3Er(pyMeO)_4$}, and {$Co_3Tm(pyOMe)_4$} optimized with the same spin multiplicity (see Tables A.9 and A.10). However, a low spin configuration on Co1 has been found to be energetically favored in the case of {$Co_3Er(pyMeO)_4$}, which is not the case for {$Co_3Tm(pyOMe)_4$} (vide infra; see Table A.8). For the S2 state, an intermediate spin state is energetically preferred where the Co1–O3 bond is elongated by 0.3 Å compared to the low spin state. The elongation of the bond results in complete breakage as can be seen from the reduction of the SEN from σ_{Co1-O3} = 0.13 to σ_{Co1-O3} = 0.00 (see Table A.11). This further leads to a more pro-

nounced change of the coordination geometry at Co1. The S2 state plays a key role in the catalytic cycle since the oxygen–oxygen bond formation by a WNA is supposed to take place there. As for many other cobalt based WOCs[251,286,287] this reaction is assumed to be rate limiting.[67,131] Comparing the (canonical) frontier orbitals of the S2 structures of $\{Co_3Ln(pyMeO)_4\}$ we find that in case of the open structure the LUMOs+X (X = 0,1,..,4), which are supposed to govern the nucleophilic attack of a water molecule,[67,131] possess significant contributions of the π^*_{Co-O} orbitals and the π^*_{hmp} orbitals (see Figure A.5, A.4, A.6, and A.7). How the admixing of the π^*_{hmp} orbitals to the π^*_{Co-O} orbital could benefit a WNA is not obvious. The most striking difference between the open and closed structure is the stability of the relevant LUMOs. Those orbitals are significantly destabilized in the open form whereas the closed one features energetically low lying accepting orbitals, which in general facilitate the nucleophilic water attack.[67,131,175] The reactivity of the S2 state is further characterized by the oxo-oxyl resonance,[288] which in turn is defined by the bond order, and therefore the SEN.[273,289,290] In the open structure the SEN between Co1 and its oxo-ligand (σ_{Co-O} : 0.43) is reduced by 21% compared to the closed one (σ_{Co-O} : 0.54), which is in accordance with an elongation of the bond by 0.07 Å. The stronger radical character should potentially enhance the reactivity of the ligand and thereby facilitate the oxygen-oxygen bond formation.[119] However for a conceptually similar Mn-oxo system, the radical character alone was found not to be crucial for a WNA.[26]

The energetics of the catalytic intermediates after the WNA shows that the closed structures are always thermodynamically favored compared to the open ones. From a structural point of view those intermediates are again very similar for both catalysts, except for the orientations of the hydroperoxo and hydroxyl ligands. The increased stability of the closed model might (at least partly) originate from an intramolecular hydrogen bond formed between the hydroperoxo ligand and a hydroxyl ligand of the active center (see Figures 4.2, A.3 and A.8). This, though, may have been favored due to the use of the implicit solvent model COSMO, which has shortcomings with respect to proper inclusion of hydrogen bonding effects.[291]

The high spin state of the Co1 metal center is energetically favored for all open cubane structures except for the S2 state where the low spin state (in the case of $\{Co_3Er(pyMeO)_4\}$) and an intermediate spin state (for $\{Co_3Tm(pyOMe)_4\}$) are preferred (see Tables A.6 and A.7). However, the different spin states are often within an electronic energy range of less than 5 kcal mol^{-1}. This is also true for the closed structures (except for the S2 state where the intermediate spin multiplicity leads to a significantly lower electronic energy; see Table A.5). Those findings suggest that multiple spin states are accessible for each intermediate and energetically unfavorable spin-crossing events can be avoided.

The free energy differences between the catalytic intermediates and their respective ground state are given in Figure 4.3. While for the open and closed structures are similar for the first intermediate (S1), we find that the difference becomes larger along the catalytic pathway. Those findings bring to mind the relative stability of the intermediates (see Figure 4.2) and indicate that the open structures do not improve the thermodynamics. This is true for both $\{Co_3Er(pyOMe)_4\}$ and $\{Co_3Tm(pyOMe)_4\}$, since the free energy differences rather increase than decrease compared to the ones of the closed structures.

This raises the question whether a change from open to close or vice versa is fea-

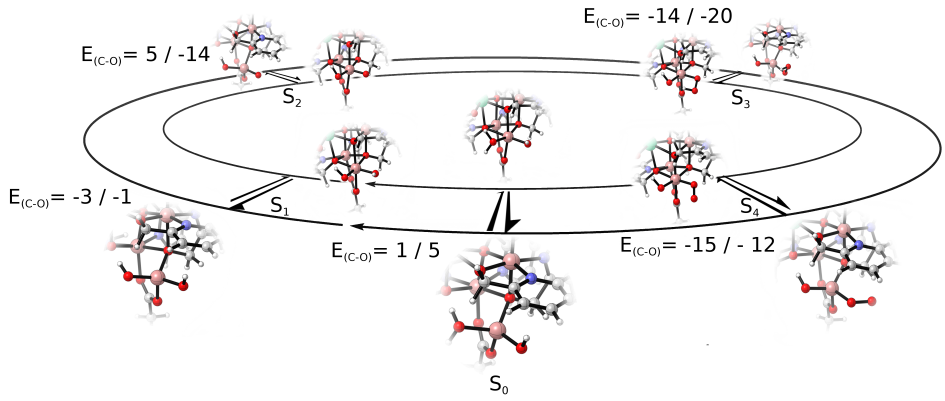


Figure 4.2: Schematic representation of the catalytic water oxidation mechanism: The inner cycle represents the one based on the previously studied closed cubane core structures[14] while the outer hosts the catalytic intermediates with an open cubane core. The difference in zero-point-corrected electronic energy between the two conformers is given as in kcal mol^{-1}(calculated with COSMO). The numbers on the left hand side of the slash correspond to for $\{Co_3Er(pyMeO)_4\}$, the ones on the right hand side to $\{Co_3Tm(pyOMe)_4\}$. For the sake of clarity only the active center (Co1) and its ligands are shown (see Figures A.2 and A.3).

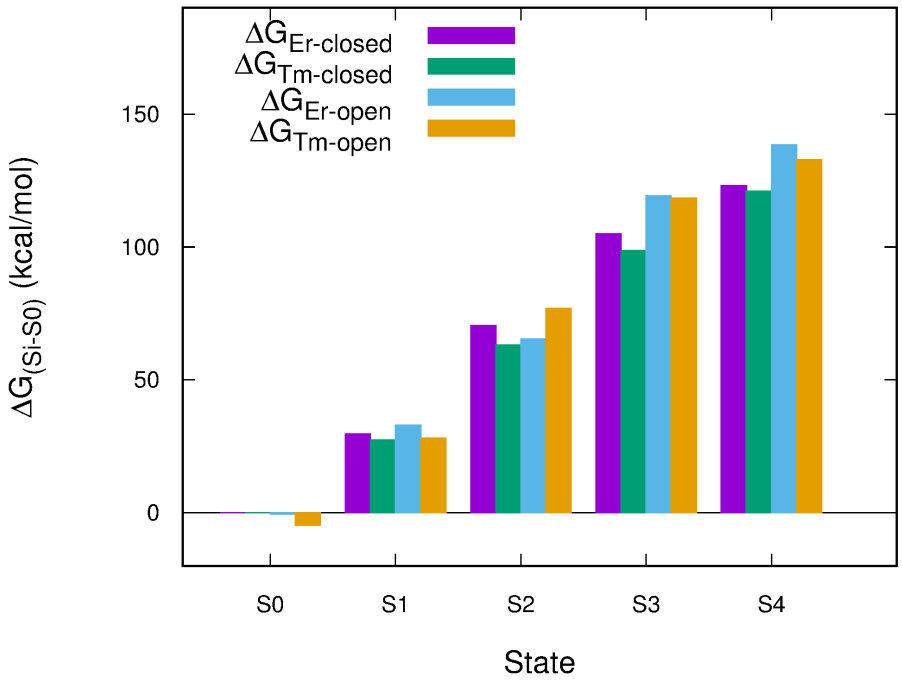


Figure 4.3: Differences in free energy (ΔG_{Si-S0}) calculated for $\{Co_3Er(pyOMe)_4\}$ and $\{Co_3Tm(pyOMe)_4\}$ relative to their catalytic ground state (S0).

sible under catalytic conditions. Therefore the energetic barrier for such a structural amendment was calculated in the catalytic ground state (S0) (see Section 4.1 for computational details). A possible minimum electronic reaction path for the transition from the closed to the open cubane motif is shown in Figure 4.4, the electronic energy barrier is estimated to be 6 kcal mol^{-1} for both $\{Co_3Er(pyOMe)_4\}$ and $\{Co_3Tm(pyOMe)_4\}$. The reaction coordinate of both catalysts is highly complex involving multiple bonds and angles. In particular the Co1–O1, and Co1–O3 bonds are elongated during the transition (see Tables A.9 and A.10). Further the intramolecular hydrogen bond between the aqua-ligand of Co1 and the hydroxyl of the lanthanide is broken and replaced by a hydrogen bond with O1 (for more details, see Figure A.1).

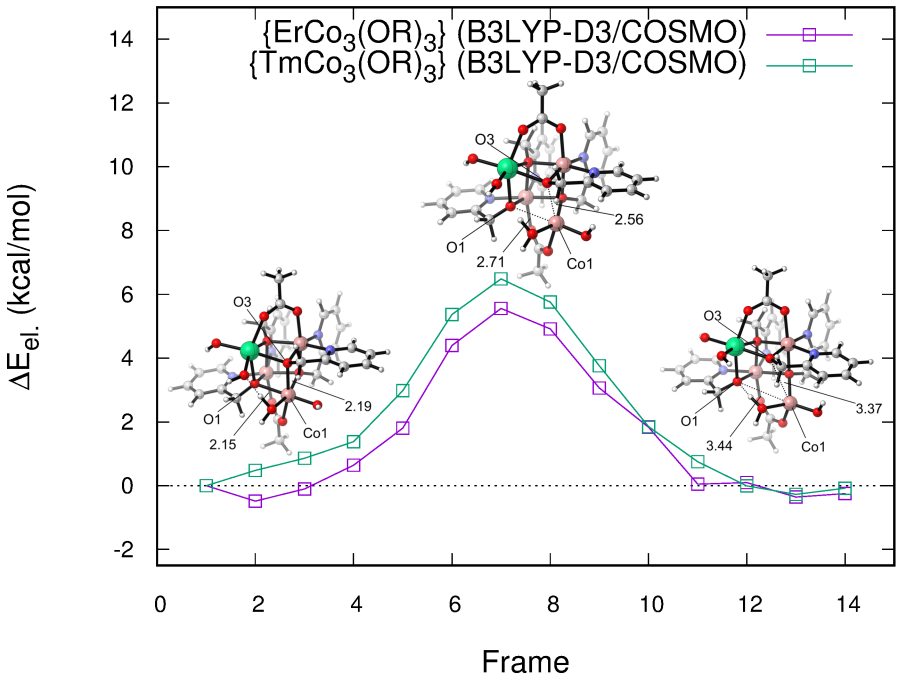


Figure 4.4: A possible reaction path for the transition from the closed to the open structure in the catalytic ground state (S0) is visualized for both $\{Co_3Er(pyOMe)_4\}$ and $\{Co_3Tm(pyOMe)_4\}$. The Co1–O1 and Co1–O3 bond lengths are given in Å.

Since the structural change is comparable for all catalytic states, we assume the kinetic barriers to be similar as well. The equilibrium between the two structures is thus mainly driven by their thermodynamics (see Figure 4.2). This suggests that 'mixed' cycles composed of both open and closed cubane core structures (see Figure 4.5) could exist. From the relative stabilities shown in Figure 4.2 it becomes evident that the open motif may only contribute significantly in the ground state (S0) and the first two intermediates (S1 and S2). In case of $\{Co_3Er(pyOMe)_4\}$ there are two 'mixed' water oxidation pathways (see Tables A.12 and A.13 for all reasonable pathways) which either resemble the thermodynamics of an ideal catalyst as good as an all-closed pathway or even slightly better (see Figure 4.5). A common feature of both 'mixed' pathways is the open structure of the catalytic ground state, which however is only slightly more stable than the closed one (about 1 kcal mol^{-1}). If the first in-

termediate also remains in its open form, the pathway becomes slightly unfavorable by 3 kcal mol^{-1}, while the open structure in the S2 state again decreases the free energy difference by 5 kcal mol^{-1}. If we take all possible combinatorial pathways into account, an open-closed-open-closed-closed pathway is found to resemble an ideal catalyst best. However, such a pathway would require four transitions from open to close cubane core (or vice versa) which – even though the barriers are expected to be reasonably small – is presumably rather unlikely. For $\{Co_3Tm(pyOMe)_4\}$ there is no mixed pathway that is energetically more favorable than the all-closed one. However if solely the first intermediate obtains an open cubane core structure the free energy differences increase by only 5 kcal mol^{-1}. In summary the all-closed pathway discussed in detail in by Hodel and Luber is still the most likely one from a thermodynamic point of view.[67] Nevertheless, our findings strongly suggest that at least for the catalytic ground state (S0) both conformations can coexist.

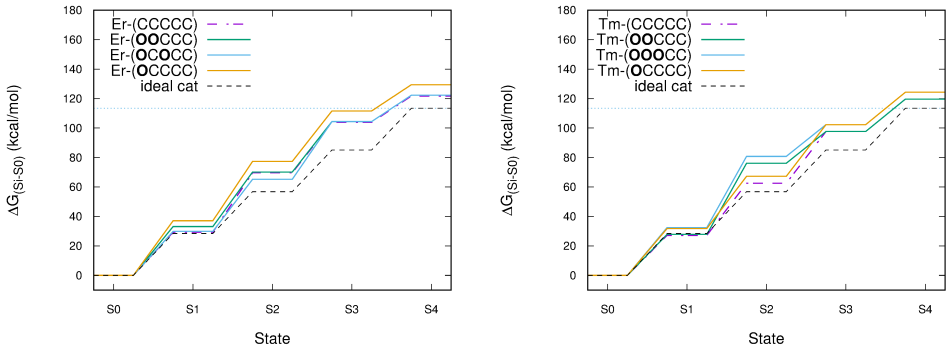


Figure 4.5: Relative differences in free energy for catalytic cycles composed of O(open) and C(closed) conformers; The horizontal line represents the thermodynamic limit (i.e. the experimental value of 113.5 kcal mol^{-1} at pH 0 required for the oxidation of water). Reaction pathways obtained for $\{Co_3Er(pyOMe)_4\}$(left); Reaction pathways calculated for $\{Co_3Tm(pyOMe)_4\}$(right).

4.2.3 Solvation Effects

In the following section we will address the effects of the solvation model on the different structures.[68] An inherent problem of many implicit solvent models (such as COSMO) is their inability to accurately describe directed bonds such as hydrogen bonds between solute and solvent. In order to attenuate this limitation, we additionally carried out DCOSMO-RS calculations (see Section 4.1).[291] In general the thermodynamics of the open and closed forms are only slightly affected by the use of DCOSMO-RS. Except for the state S2, free energy differences obtained using COSMO and DCOSMO-RS are within 5 kcal mol^{-1} (see Tables A.4 and A.5). The improved treatment of hydrogen bonding by DCOSMO-RS results in a stabilization of the S2 state by 12 kcal mol^{-1} in the case of $\{Co_3Tm(pyOMe)_4\}$ and 9 kcal mol^{-1} for $\{Co_3Er(pyOMe)_4\}$, respectively. This is in particular important for the oxo-ligand, which mainly distinguishes the S2 structure from the previous intermediates. The strength of the solute-solvent interaction might be characterized by the surface charge density obtained from COSMO / DCOSMO-RS calculations (see Figure 4.6). In the case of the open cubane structure, a decrease of the surface charge density at the oxo-ligand is observed when DCOSMO-RS is used instead of COSMO,

suggesting weaker interaction between the solvent and the ligand, while for the closed cubane structure the surface charge density is increased at the oxo-ligand implying stronger interactions with the solvent. Relating the weaker solute-solvent interactions in case of the open cubane structure to the oxyl-character of the ligand reveals a reduced hydrophilicity of the oxyl-ligand. Recently it has been reported that the oxyl-ligand of Ru^{V}(bda) (bda : 2,2′-bipyridine-6,6-dicarboxylate) is not susceptible towards hydrogen bonding.[292] Those findings are in good agreement with the results from the DCOSMO-RS calculations on the S2 state of $\{Co_3Ln(pyMeO)_4\}$.

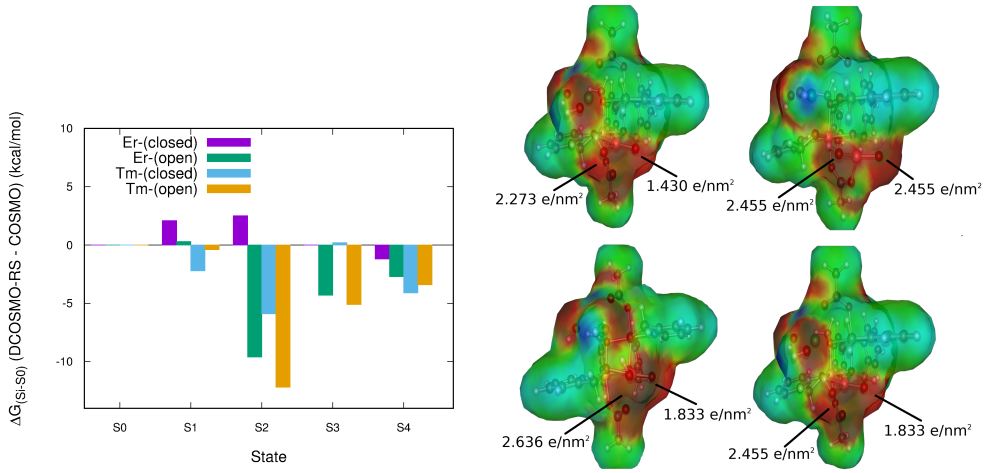


Figure 4.6: Difference in free energy differences (ΔG_{Si-S0}) obtained employing DCOSMO-RS or COSMO for catalytic water oxidation by $\{Co_3Er(pyOMe)_4\}$ and $\{Co_3Tm(pyOMe)_4\}$ (left). Surface charge density maps of intermediate S2 ($\{Co_3Tm(pyOMe)_4\}$) (right), obtained by using COSMO (top: closed (left), open (right)) and DCOSMO-RS (bottom: closed (left), open (right)).

Going beyond solvent continuum models in analogy to previous work,[67,68,131] the structures were also modeled with an explicit first solvation shell derived from a DFT-based MD run (see Section 4.1 for details). A detailed description of the structures can be found in the Appendix in the Section A.7. Here, we will only address the important features already discussed in the context of the implicit solvation model. As mentioned earlier, a proper inclusion of solvent-solute interactions is in particular important for the S2 state. When employing explicit solvation, oxo-H distances of 1.7 Å were found. Such distances are within the range of hydrogen bonds between solvent molecules which are in general between 1.5 - 1.9 Å. The hydrophilic behavior of the closed cubane core oxo-ligand was already discussed by Hodel and Luber,[67] where prior to the WNA a PT from the attacking water to another solvent molecule and finally to a hydroxyl ligand of the lanthanide center was observed. The apparently different behavior for conceptually similar catalysts[292] can easily be resolved since the Co^{III}-oxyl species possess two additional electrons compared to Ru^{IV}-oxyl, which in term increase the basicity of the oxyl-ligand and therefore favor hydrogen bonding. Another consequence of the two electrons is the reduced electrophilicity of the Co^{III}-oxyl that makes a WNA less likely.[67] Those findings are also in agreement with the COSMO and D-COSMO calculations where the oxyl-ligand was found to be hydrophilic. Another important feature which might be described more accurately using explicit solvation is the formation of intramolecu-

lar hydrogen bonds as observed for the state S3. Such interactions are only observed if the active center (Co_1) is in the high-spin state. From an energetic point of view such a high spin configuration is only favorable for $\{Co_3Tm(pyOMe)_4\}$ but not for $\{Co_3Er(pyOMe)_4\}$. A simplistic interpretation of the correlation between the preferred spin state and the structure of the catalyst is not possible, since the explicit solvation shell has a considerable influence on the orientation of the ligands. Besides structural features, the explicit solvation also influences the stability of the open/closed forms. This is measured by the free energy differences (see Tables A.1 and A.2). For $\{Co_3Er(pyOMe)_4\}$, the open form of the states S0 to S2 are stabilized by 18 kcal mol^{-1} - 26 kcal mol^{-1} compared to the closed ones, while for the S3 and S4 states the closed model is energetically favored. For $\{Co_3Tm(pyOMe)_4\}$, there is an increasing stabilization of the open structure during the catalytic cycle, starting from 6 kcal mol^{-1} at the S0 state and reaching almost 30 kcal/mol in the S4 state. Comparing the relative stabilities of the two structural models obtained with implicit and explicit solvation (see Figure 4.2 and see Tables A.1 and A.2), we find the open structures to be stabilized by the solvent molecules by 15 kcal mol^{-1} - 25 kcal mol^{-1} in case of $\{Co_3Er(pyOMe)_4\}$, and by 2 kcal mol^{-1} - 40 kcal mol^{-1} in the case of $\{Co_3Tm(pyOMe)_4\}$, respectively. However, it should be kept in mind that the calculation with explicit solvation shell has been carried out within the CP2K package whereas Turbomole was used for the COSMO calculations using different settings (e.g. basis set), which lead to additional differences in the electronic energies. To verify whether the stabilization originates primarily from the hydrogen bonding network within the solvent or from the opening of the cubane cage, we recalculated the electronic energy differences of the cubane structure without solvent molecules (within CP2K). From the absolute differences in free energy it becomes evident that for both catalysts the open structures of the intermediates of the S0 and S1 states are more stable than the closed ones (see Tables A.1 and A.2). For the other intermediates, the open and closed forms have either a similar stability or the closed one is energetically favored. There is no consistent trend how the explicit solvation affects the stability of the intermediates in the S2 and S3 states. The lack of a consistent trend might points towards a major disadvantage of a static explicit solvation model. The existence of many shallow local minima on the potential energy surface makes the comparison of energies a tedious task since one has to assure to compare the same local minima with respect to the solvation shell. To get a more accurate description of explicit solvation approaches, sampling methods such as DFT-based MD might be used as described elsewhere.[68] Nevertheless, such methods are significantly more expensive in terms of computational cost and beyond the scope of this work.

4.2.4 Opening of the Cubane-Core

Further, we calculated the barriers for the opening of the $\{Co_3Ln(pyOMe)_4\}$ cubane core in the catalytic ground state (S0) using NEB calculations. Most of the solvent molecules are barley affected by the structural change of the catalyst, only the few molecules close to the active center (Co1) are slightly pushed away from their initial positions due to the opening of the cage. For both catalysts $\{Co_3Er(pyOMe)_4\}$ and $\{Co_3Tm(pyOMe)_4\}$ electronic barriers of 3 kcal mol^{-1} were found. Those barriers are slightly smaller than the ones obtained when employing the implicit solvation model, and emphasize that both the open and closed form of the catalytic ground state can co-exist (see Figure A.9). To compare the relative stability of

the intermediates optimized with implicit or explicit. solvation directly, electronic energies of the explicitly solvated models were recalculated within TURBOMOLE (with COSMO), where the solvent molecules were deleted. As expected all non-optimized structures are higher in electronic energy than the optimized ones (see Table A.3). However, the free energy differences are for some intermediates as small as 0.3 kcal mol^{-1}. This clearly suggests that the open and closed form could co-exist for $\{Co_3Er(pyOMe)_4\}$. In case of $\{Co_3Tm(pyOMe)_4\}$, the explicitly solvated structures are significantly more destabilized compared to the optimized ones by 12 kcal mol^{-1} - 30 kcal mol^{-1}. The preferred spin state strongly depends on the solvation model and the used computational settings. For the explicit solvation model (CP2K) (see Tables A.6 and A.7), the high spin states are favored for the open model in the S0 and S1 states while for the S2 to S4 states a low spin state on the Co1 metal center is preferred. When the solvent molecules were deleted the trend for the later intermediates inverts (see Table A.8), and we find the intermediates with a high-spin configuration to be energetically favorable. There is a good agreement between the preferred spin states on Co1 for the open structures calculated in TURBOMOLE (with COSMO) and the CP2K structures without explicit solvation, recalculated in TURBOMOLE (with COSMO). Nevertheless, there are no obvious trends with respect to certain spin states for both the two motifs and the two catalysts. The distinct behavior of $\{Co_3Er(pyMeO)_4\}$ and $\{Co_3Tm(pyOMe)_4\}$ is related to structural differences (see Hodel and Luber for a detailed description).[67] The solvation model chosen affects the nuclear and electronic structure, which can change the energetic ordering of the spin states. A more accurate approach beyond DFT would rely on multireference methods, whose application to the highly complex electronic structure of the cubanes presents currently a major challenge, in particular if also solvent effects should be considered.

4.3 Summary and Conclusion

In summary, we have presented novel structures of the recently presented bioinspired $\{Co_3Ln(pyMeO)_4\}$ (Ln = Er, Tm) WOCs. Their cubane core can open giving access to a so-called 'open' motif. At least for the first two intermediates of the proposed catalytic cycle based on nucleophilic water attack, the two forms (open and closed) are similar from a thermodynamic point of view and kinetically accessible. This further suggests that the 'open' cube structures actively contribute towards water oxidation, which is to some extent reminiscent of recent discoveries regarding the OEC.[132] Comparison of the 'open', 'closed' or 'mixed' catalytic cycle with the one of a thermodynamically ideal catalyst reveals that in general the 'closed' structures better resemble such an ideal catalyst. However, the increased versatility of the coordination geometry of Co1, which may change from octahedral to tetrahedral and vice versa during the water oxidation process, highlights the fact that the ligand environment for those homogeneous catalysts is not completely rigid and therefor might play a crucial role in terms of catalytic activity and long term stability.

The unexpected discovery of the novel structure encouraged us to have a closer look at the influence of the solvation model. When using an explicit solvation shell without applying sampling techniques such as DFT-MD, care has to be taken that the solvation shells resemble the desired local minima. Nevertheless, with an improved description of important solute-solvent interactions structures may be found which would not be detected otherwise. The open cubane core structures initially were

only discovered using an explicit solvation model as well. We were also able to prove that they are minima on the potential energy surface using an implicit solvation model. In particular the hydrogen bonding of the solvent and solute appears to play a crucial role in stabilizing certain conformations. Especially the solvation of the oxo-ligand in the S2 state as well as the one of the hydroperoxo ligand in the S3 state have been found to be affected by the solvation model. Employing DCOSMO-RS revealed a significant difference in the description of the oxo-ligand in the 'open' and 'closed' form compared to the one obtained using COSMO, which might be related to radical character of the oxo-ligand. The hydrophilicity of this species is expected to directly affect the likelihood for a WNA and is therefore considered to be a key aspect in modeling water oxidation reactions.

Chapter 5

Towards the Rational Design of the Py5–Ligand Framework for Ruthenium–Based Water Oxidation Catalysts

This chapter is a shorted version of: Schilling, M.; Böhler, M.; Luber, S.; *Dalton Trans.* **2018** *47*, 10480–10490.[174] If not stated otherwise all figures of this chapter were published in the reference mentioned beforehand.

Together with our experimental collaborators we have studied a series of mononuclear ruthenium-based WOCs which possess an intramolecular base in form of a dangling pyridine.[173,293] The ligand-framework of those catalysts (Py5R) is composed of two bipys as well as a py unit liked together at an sp^3 carbon. (see Figure 5.1). Two derivatives of the WOCs that were only different in the fourth substituent of the bridging carbon (either a methyl or a methoxy group) were studied by DFT simulations. Despite the different catalytic activities observed in experiments we found that both the thermodynamics to reach the catalytically active Ru^V=O as well as the kinetics of the WNA are similar for both catalysts.[173] Inspired by this curiosity, we further investigated those catalysts.

In this chapter we focus our attention in particular on the ligand framework and how it affects the catalytic performance (a brief summary of the key findings of this chapter is given in Section 1.4.4). The idea behind this was not only to further elucidate experimental findings but also to deduce design guidelines to further improve the catalysts. This computational work is structured as follows: First, we will discuss the Py5–ligand framework, its coordination mode and the resulting isomers. In the second part we focus on the water oxidation mechanism, and in particular on the O−O bond formation by a WNA. Finally, in the last part we explore modifications of the ligand framework and how they affect different parts of the catalytic cycle.

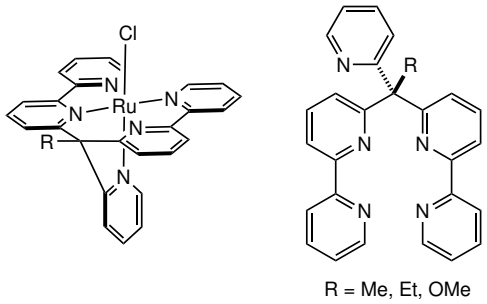


Figure 5.1: Schematic drawing of the Py5 derived Ru-based WOCs studied in this work.

5.1 Computational Methodology

All Calculations were carried out at the B3LYP-D3/def2-TZVP(COSMO)//BP86-D3/def2-TZVP level of theory (see Section 2.5.1 for more details) using the Turbomole software package (version 7.01)[206].

5.1.1 Thermodynamic Properties

Deprotonations include the solvation free energy of a proton (but not its translational free energy) of 265.9 kcal mol^{-1}; reduction potentials are given relative to the standard hydrogen electrode (SHE) which was taken to be 4.28 V.[76,294,295] The standard reduction potential was calculated as given below, where n is the number of electrons transferred, F is the Faraday constant, $\Delta G^{\circ}_{\text{reduction}}$ the Gibbs free energy change of the reduction and $E^{\text{absolute}}_{\text{SHE}}$ the absolute reduction potential of the SHE:

$$\Delta E^{\circ}_{\text{reduction}} = \frac{\Delta G^{\circ}_{\text{reduction}}}{n\text{F}} - E^{\text{absolute}}_{\text{SHE}}. \tag{5.1}$$

PCET reactions include the energy of half a hydrogen molecule[268] whose Gibbs free energy correction consists of the zero-point vibrational energy plus a thermal correction of 0.52 eV.[269] For comparison with the ideal catalyst (and only there), PCET free energies were scaled by 0.807 which is the ratio of the computational and the experimental free energy for water oxidation with our settings. Larger basis sets were used to calculate single-point electronic energies (def2-TZVPP, def2-QZVP) but did not influence the relative energies significantly. Electronic contributions of the multiplicity on the ground state was neglected when computing the Gibbs free energy as it would amount to only 0.41 kcal mol^{-1} for doublet states and 0.65 kcal mol^{-1} for triplet states.

5.1.2 Bond Scans

The scans were done with geometries optimized at the BP86-D3/def2-SVP level and B3LYP-D3/def2-TZVP single-points. A good nuclear structure can usually be obtained already on the def2-SVP level and for our scans the reduced computational time was crucial. All scans were calculated as restricted singlets on the Ru$^{\text{II}}$ oxidation state and a step-width of 0.1 Bohr was used around the equilibrium geometry (3.3 to 5.0 Bohr) and 0.2 Bohr for the rest. For the electronic energy calculations

(in the def2-TZVP basis) the proper propagation of the orbital guess for the self-consistent field procedure and damping of proved thereof be critical in order to obtain smooth curves.

If not explicitly stated otherwise, all reported energies are Gibbs free energies in $\mathrm{kcal\,mol^{-1}}$.

5.2 Results and Discussion

As laid out in the introduction, the objective of this work is to get a better understanding of the Py5 ligand framework, in order to derive design guidelines on how to improve Py5 containing WOCs.

5.2.1 Water Association

From experiments it is known that the Py5 ligand framework can arrange in a *cis* or *trans* fashion (relative orientation of bipy fragments) where by the *trans*-bipy isomer was found to be the dominant one.[172,173] In our previous work we have established that the differences in catalytic efficiency (TON/TOF) between the Py5Me and Py5OMe ligand framework are unlikely to be caused by the thermodynamic energy differences on the path to the catalytically active Ru^{V}-oxo or the barrier for the WNA afterwards.

The focus of this work is therefore primarily on the initial steps of the catalytic mechanism, namely the association of a water ligand.[173] This is a prerequisite for water oxidation since the catalyst itself does not bear an aqua ligand. For similar ligand frameworks Sun and co-workers were able to isolate intermediates with a sevenfold coordinated Ru–center.[296] Such an extension of the coordination sphere makes a ligand dissociation prior to water association unnecessary. However, we (computationally) did not find any stable sevenfold coordination for catalysts derived from the Py5 ligand framework. To the best of our knowledge, there is no experimental evidence for a sevenfold coordinated intermediate, therefore we assume that a ligand exchange reaction has to take place. What is more, there is neither an experimental clue nor evidence for a sevenfold coordinated intermediate, therefore, we can safely assume that a ligand exchange reaction needs to take place.
In our previous work we naturally assumed that the pyridine fragment is displaced by a water molecule resulting in a *trans* alignment (with respect to the chlorido ligand).[173] This option will be discussed later in detail in Section 5.2.2. An alternative pathway involves the displacement of a pyridine subunit of a bipyridyl fragment, the resulting isomers are referred to as *cis*. These *cis* structures were found to be 8 to 20 $\mathrm{kcal\,mol^{-1}}$ higher in energy than their respective *trans* counterparts, which in turn makes their involvement in the catalytic cycle unlikely (see Figure B.1). This supports our previous conclusion that the *trans* isomer is the most likely candidate to undergo water oxidation. Nevertheless, there are two interesting points to note. Firstly, it appears that these *cis* structures are generally slightly more acidic. As can be inferred from the more frequent intramolecular PTs during geometry optimizations, which lead to a release of sterical strain in the structure. Secondly, the *cis*-Ru^{V}=O structure fundamentally differs from *trans*-Ru^{V}=O in that the oxygen has been subjected to a nucleophilic attack by a pyridine fragment giving rise to a formal N–O bond.

The last option is a water-chlorido exchange. Those structures were also found to be too high in energy in order to contribute to water oxidation catalysis, in part due to their higher charge which renders oxidation reactions energetically more demanding. Furthermore, those *trans*-aqua-pyridine intermediates are the precursors of bis-aqua isomers. Those have been found to possess a significantly lower catalytic activity than their chlorido analogues.[173] This led to the speculation that the bis-aqua isomers are part of a deactivation pathway which results either in inert side–products ore even in the decomposition of the catalyst.
Therefore we reached the conclusion that a strong Ru – Cl bond is a key requirement to maintain the high catalytic activity of Py5 derived WOCs. This is rather counter-intuitive since for similar ruthenium based WOCs, halides such as iodo ligands were found to be especially labile towards ligand exchange reactions.[297] Note, Py5 bearing WOCs do require a halide as spectator ligand, while for many other WOCs the halides are mere placeholders for the substrate to coordinate the metal center. The experimental rates for the chlorido-water exchange were found to be one order of magnitude larger for Py5Me compared to Py5OMe.[173] Those differences can not be fully explained based on the thermodynamics which are similar for both ligands. For a more complete picture, the concerted TSs of the chlorido-water exchange reaction were modeled employing the same computational procedure as for the WNA-TS (see Figure 5.2). The obtained TS energies range from 21 to 25 kcal mol^{-1} for Ru^{II} and 25 to 29 kcal mol^{-1} for Ru^{III}, respectively (see Table 5.1). Unlike the thermodynamic which suggests that there is no significant difference between the two ligands Py5OMe and Py5Me, the activation barriers tell a different story. The barriers for Py5Me are independent of the oxidation state about 4 kcal mol^{-1} smaller than the respective Py5OMe barriers. This observation allows finally to rationalize the above mentioned experimentally observed differences.[173]

Variant	TS
Ru^{II}Py5OMe	25.3
Ru^{III}Py5OMe	29.0
Ru^{II}Py5Me	20.9
Ru^{III}Py5Me	25.1
Ru^{III}Py5OMe-*p*OMe	28.8
Ru^{III}Py5OMe-*p*NO_2	29.2

Table 5.1: Activation energies (kcal mol^{-1}) of the concerted chlorido-water exchange, relative to the energy of the respective chlorido-species and a water molecule. Energies were obtained at the B3LYP-D3/def2-TZVP/COSMO level of theory.

Additionally, the dissociative pathway was modeled for the Ru^{III} species, where energy differences between the two ligand frameworks were found to be as close as 1.6 kcal mol^{-1}. These intermediates are approximately 10 kcal mol^{-1} higher in energy than the concerted TS, which renders them chemically unimportant. In a similar study the ligand exchange reaction on Ru-aqua ions were investigated, where both the dissociative and the interchange (concerted) mechanism were found to be feasible depending on the oxidation state of the metal center.[298] This might be explained by the rigidity of the Py5 ligand framework which cannot adopt to the under-coordinated metal center in the dissociative pathway as easily as aqua

ligands. This leaves the concerted mechanism as the sole feasible pathway for a chlorido-water exchange.

Having identified a key difference between the two ligands we were interested whether the Ru−Cl bond could be altered by introducing substituents at the pyridine fragment *trans* to the halide. Therefore, activation barriers were calculated for catalysts, where a nitro or methoxy group was introduced in *para* position of the pyridine fragment in order to electronically alter the N−Ru bond. Different substitution patterns are possible, however in order to keep electronic and steric effects separated substitutions in the *para* position are presented. However, those changes had a minor influence on the energetics of the TSs (see Figure 5.2).

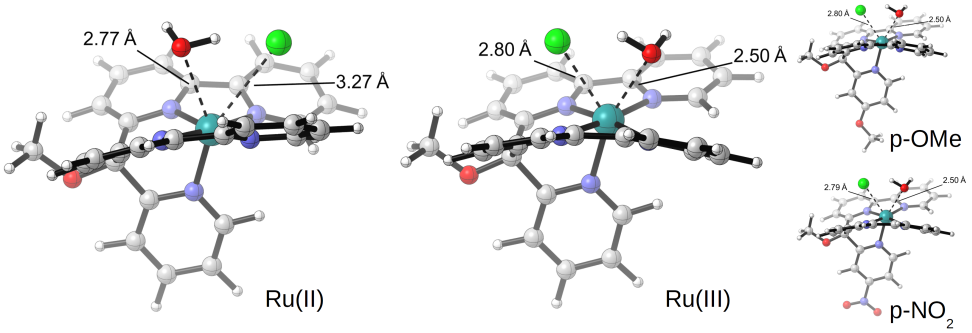


Figure 5.2: TSs for chloride displacement in Py5OMe with Ruthenium in oxidation state II or III assuming that the process is concerted. Note that the positions of chloride and water are interchanged for Ru^{II} and Ru^{III}, this is not a computational result but a consequence of different starting guesses. Structures were were optimized at the BP86-D3/def2-TZVP level of theory.

5.2.2 Water Oxidation Mechanism

Having established which isomers contribute to the catalytic performance, we now devote our attention to the thermodynamics and kinetics of the water oxidation mechanism itself.

Water Activation

In our previous study we established that the thermodynamics of the water-oxidation cycles up to the Ru^{V} oxidation state are virtually identical for both ligands Py5OMe and Py5Me. This is with the exception of the substitution of pyridine ligand with a water molecule, which differs significantly.[173] Those findings indicate that the two ligands most likely distinguish themselves in terms of their electronics but mainly in the sterical demand of their substituents. To further elucidate this assumption we performed the same analysis of both thermodynamics and kinetics for the hypothetical ethyl-analogue (Py5Et) of the Py5 ligand framework, which we assumed would impose sterical restrictions more similar to those of the Py5OMe ligand, while being electronically close to Py5Me. The energies obtained with this hypothetical ligand are given in Figure 5.3. As conjectured, they are in essence identical with the Py5OMe and Py5Me containing complexes, except for the pyridine-water substitution where the energies lay approximately in between the other two ligands (see Table B.1).

Adding to the discussion of the previous section, the energies calculated for the thermodynamics of the chlorido-exchange of Py5Et are in between the ones of Py5OMe, and Py5Me, which highlights the similarity of those ligands (see Table B.1). It also points out that, at least for the thermodynamics, sterics plays a negligible role. On the other hand, sterics might play a crucial role in the water-association (see Figure 5.4). In the Py5OMe ligand the interactions between the oxygen and the bipyridyl fragments are negligible compared to the strong methyl–bipy interactions in the case of Py5Me, which originate from the proximity of the hydrogen atoms to the bipyridyl fragments. From this observation, differences in terms of thermodynamics and kinetics might be expected in reactions involving the mentioned site. However, experimentally the rates for pyridine-water exchange were found to be virtually identical.[173] The discrepancy between theory and experiments might be manifold, starting form the simplified computational model system which neglects any kind of hydrogen bonding with the solvent, ending at the experimental techniques and fitting procedures, which are all prone to uncertainties.

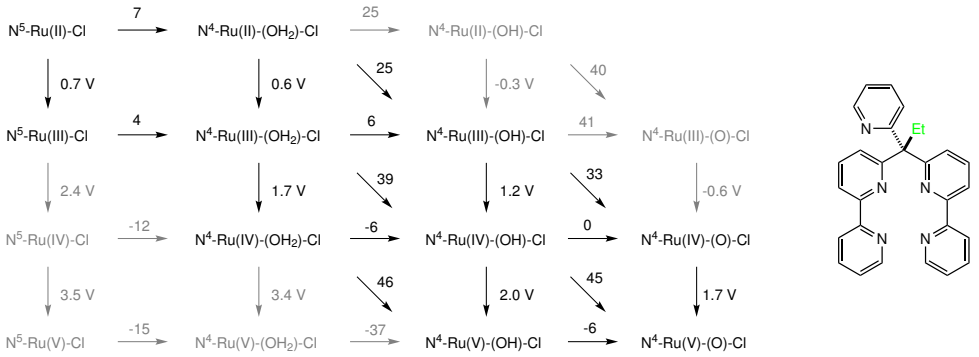


Figure 5.3: Thermodynamic table up to $Ru^{V}=O$ for Py5Et containing complex. Horizontal arrows represent water association (first) or deprotonations (in kcal mol^{-1}), diagonal arrows stand for PCETs (in kcal mol^{-1}), vertical ones represent reduction potentials given in V. Note that $Et-N^{4}-Ru^{IV}-(OH_2)-Cl$, $Et-N^{4}-Ru^{V}-(OH_2)-Cl$ and $Et-N^{4}-Ru^{V}-(OH)-Cl$ are N-H acidic due to a PT from the oxygen to pyridine nitrogen. Structures associated with deprotonation energies larger than 20 kcal mol^{-1} and reduction potentials larger than 2.0 V have been deemed unlikely and marked in gray. Energies were obtained at the B3LYP-D3/def2-TZVP/COSMO level of theory.

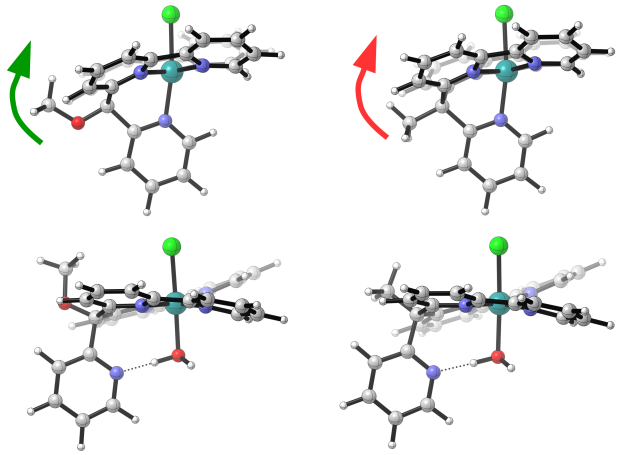


Figure 5.4: Illustration of the sterical interaction which occurs upon pyridine dissociation for the Py5OMe ligand (left column) and the Py5Me ligand (right column). Note how the methoxy group gives lower dissociation energies because the sterical bulk is further away and does not interact as strongly with the bipyridine framework as the methyl substituent does. The same is true for the ethyl substituent.

Oxygen-Oxygen Bond Formation

Since the ethyl substitution did not significantly alter the thermodynamics, we further investigated its influence on the kinetics of the WNA, by modeling the relevant TSs. A key requirement for modelling these TS states is the inclusion of an additional explicit solvent molecule which forms a hydrogen bond with the nucleophile (see Figure 5.5). The necessity to include explicit solvent molecules for WNA TSs was also observed by others.[34,35,299] For a more detailed validation of the explicit solvation model for the TS, we refer to Section B.2 in the appendix and in particular to Table B.2 and Figure B.2.

Unsurprisingly the activation barrier for the hypothetical Py5Et ligand matches the ones calculated for the Py5OMe/Py5Me ligands (see Table 5.2), further supporting our conclusion that substitutions at the sp^3 carbon do not significantly affect the water oxidation reaction mechanism, neither from a thermodynamical nor from a kinetic point of view.

Variant	AR	TS	AP
Py5OMe[1]	7	16	6
Py5Me[1]	5	14	9
Py5Et	7	14	9

Table 5.2: Energies (kcal mol^{-1}) of WNA structures (AR, TS, AP) with Py5 ligand containing two explicit water molecules. Energies were obtained at the B3LYP-D3/def2-TZVP/COSMO level of theory. 1: Energies published by Gil-Sepulcre *et al.*[173]

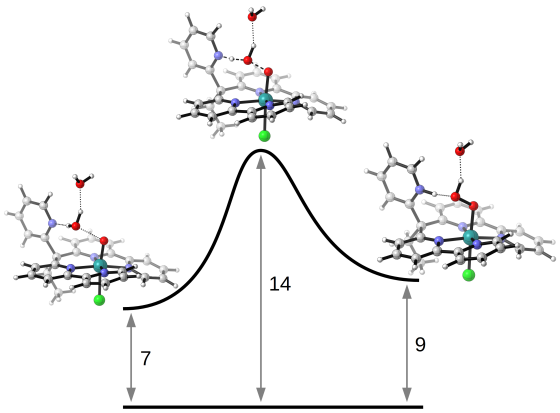


Figure 5.5: Transition-state and WNA-path for $Et-N^4-Ru^V-O-Cl$ (Py5Et). Energies given in kcal mol^{-1}. Energies were obtained at the B3LYP-D3/def2-TZVP/COSMO level of theory based on structures optimized at BP86-D3/def2-TZVP.

The Crucial Step – Replacing Pyridine by Water

In this section, we discuss the step which, besides from the chlorido-water exchange, is the most interesting one in order to differentiate between Py5OMe and Py5Me: the water association. We will refrain from further discussing the Py5Et ligand since we have shown that its energetics with respect to thermodynamics and kinetics of a WNA is in between the one for the Py5OMe and Py5Me ligands.

To further elucidate the nature of pyridine-water exchange we turned to relaxed (geometry optimized) bond dissociation scans of the $Ru-N_{pyridine}$ and $Ru-O_{water}$ bonds (see Section 5.1 for a description of the protocol). The respective starting points of the scans were chosen to be $N^5-Ru^{II}-Cl$ and $N^4-Ru^{II}-OH_2-Cl$ referred to as $Py_5OMe-Cl$ and Py_5Me-Cl, respectively. For both species, the bonds were scanned in a dissociative manner as depicted in Figure 5.6. The $Ru-N_{pyridine}$ bond dissociation is more easily achieved for Py5OMe than for Py5Me (see Figure 5.7). In contrast to this, the $Ru-OH_2$ dissociation appears to be more favorable for Py5Me than for Py5OMe (Figure 5.8). This observation mirrors the thermodynamics where the water association is much more favored for Py5OMe (easier pyridine displacement, harder dissociation of water). Furthermore, for all *trans*-aqua complexes (denoted as Py_5Me-H_2O and Py_5OMe-H_2O) dissociation of either the water ligand or the pyridine becomes more difficult due to the higher charge of the system which leads to strengthening of the $Ru-N$ and the $Ru-OH_2$ bond respectively. In the case of the *trans* aqua complexes, there is a small kink around 2.8Å, which is a computational artifact caused by the constrained $Ru-N$ bond, resulting from a rotation of the $R-C-Py$ (R = Me, OMe) fragment or a rotation along the $C-Py$ bond.

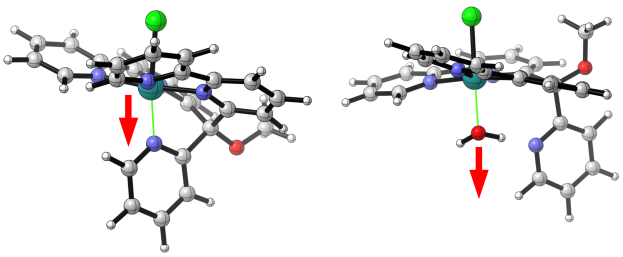


Figure 5.6: Indicated bonds in green were scanned: dissociation of Ru–$N_{pyridine}$ on the left and Ru–O_{water} on the right.

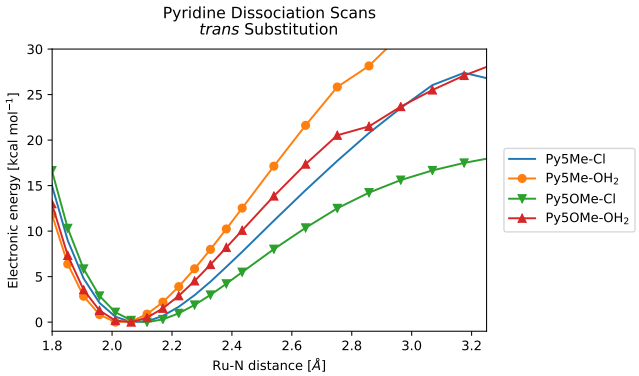


Figure 5.7: Scans of the Ru–N bond length (dissociation of pyridine), electronic energy in kcal mol^{-1}, distance in Å. Chlorido or water ligand in *trans* position to pyridine.

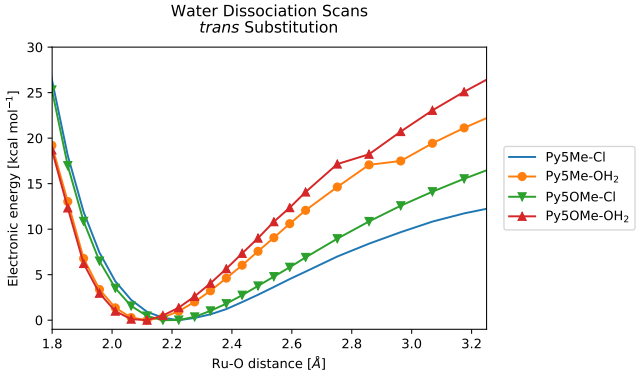


Figure 5.8: Scans of the Ru–O bond length (dissociation of water), electronic energy in kcal mol^{-1}, distance in Å.

5.2.3 Towards Ligand Design: Modification of the Axial Pyridine

As discussed in the previous section (5.2.2), the displacement of pyridine by water is probably the crucial step influencing the reactivity of the catalyst. Therefore altering the dissociation behavior of the Ru – N bond, which is governed by both sterical and electronic contributions, might be a good starting point to improve the overall catalytic performance. For the chlorido-water exchange, substitutions on the pyridine fragment turned out to have only a minor influence on the TS and its energetics (see Section 5.2.1). However for this reaction the pyridine is merely a spectator. A significant influence is expected only for reactions where the pyridine is directly involved. Combining the two concepts established in the previous sections, namely, bond-dissociation scans as well as small chemical modifications in the *ortho*, *meta* or *para* position of the axial pyridine, gives further insight into the behavior of the ligand framework.

Tuning the Water Association

First of all, the thermodynamics of the water-association reaction was investigated for a series of different modifications on the pyridine unit (see Table 5.3). Surprisingly, apart from *para*-CF_3 and *para*-CN, all substituents lead to a more exergonic water-association reaction compared to the unmodified ligand. The largest effects can be observed with substituents in *ortho* position because of their large sterical and electronic impact. The influence of *meta* and *para* substituents on the water-association is probably negligible as seen by the rather small energy differences compared to the unmodified ligand. The*ortho* substituents ordered by their decreasing influence are: (t-Bu) > CF_3 > F > Me > OMe > CN. Screening of different ligands based on thermodynamic quantities such as dissociation energies has already been successfully applied to improve WOCs (see for example the work of Duan *et al.*).[74,146,147]

Substituent	*ortho*	*meta*	*para*
H (no mod.)	4.4	4.4	4.4
OMe	3.4	4.2	4.1
F	2.8	4.1	3.9
CF_3	-1.7	4.2	5.0
CN	3.7	3.7	4.9
Me	3.0	-	-
t-Bu	-6.6	-	-

Table 5.3: Energies of the water-association reaction at the Ru^{II} oxidation state, given in kcal mol^{-1}. Modifications on the axial pyridine were introduced at the indicated positions. Energies were obtained at the B3LYP-D3/def2-TZVP/COSMO level of theory.

As for the unmodified ligand scans of the Ru – N and Ru – O bonds were performed, the former are shown in Figures 5.9 and 5.10 (the scans for the Ru – O dissociation can be found in Figures B.3 and B.4). The largest effect was observed with *ortho*-CF_3, where the larger equilibrium bond length and the flatter slope of the dis-

sociation curve suggest weaker bonding. Also fluorine, in *ortho* position, leads to a relatively small slope compared to the unmodified Py5OMe, indicating a weaker bond. However, the effect is apparently not caused by sterics since the equilibrium bond length remains the same. The potential energy curve obtained for a CH_3-substituted Py5OMe appears to be shifted towards a larger equilibrium distance, indicating some sterical interactions with the remaining ligand framework. The CF_3 modification can be viewed as a combination of the effects exerted by methyl and fluorine substituents since it shifts the bond length and displays a smaller slope than the other modifications. Consequently, all these substituents should lead to a more facile dissociation. In contrast to this, the modifications have very little impact on the water dissociation. This seems reasonable as the hydrogen-bonding capability of the pyridine is not expected to change the Ru−O bond strength significantly.

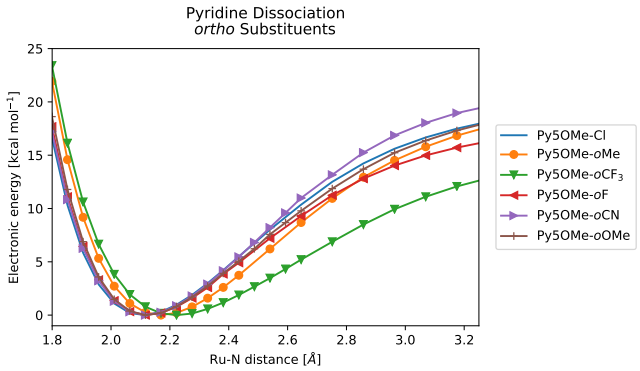


Figure 5.9: Pyridine dissociation scans, where the axial pyridine is modified by the indicated substituents in the *ortho* position, electronic energy in kcal mol^{-1}, distance in Å.

As in case of the thermodynamics, the influence of *meta* and *para* substituents on the dissociation scans shown in Figure 5.10 is minor, both for the dissociation of pyridine and water. This demonstrates the small effect that the electronic modifications of the axial pyridine in the *meta* and *para* positions have at the bonding in these molecules.

Scans performed for a higher oxidation state, i.e. Ru^{III}, gave qualitatively the same results, as the general behavior and the ordering of the substituents is conserved (see Figure B.5 and B.6).

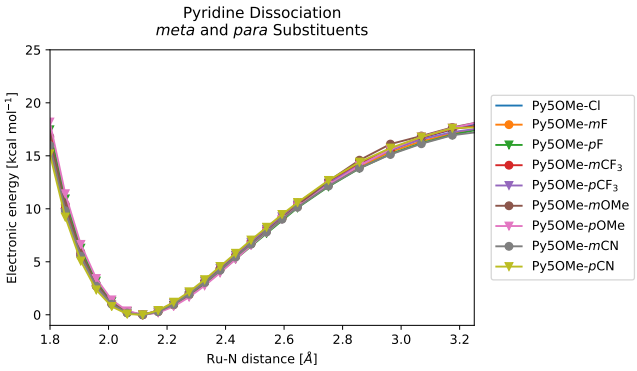


Figure 5.10: Pyridine dissociation scans with the indicated substituents in *meta* or *para* position of the axial pyridine, electronic energy in kcal mol^{-1}, distance in Å.

Influence of Modifications of the Axial Pyridine on the WNA-TS

The modifications introduced on the axial pyridine ligand naturally should not only affect the water association reaction but also the WNA. We therefore have calculated the TSs for some of the above mentioned variations (and a few additional ones). The results of those calculations are summarized in Table 5.4 which confirms that the energies are considerably affected by some of the introduced substituents. Again the largest influence is exerted by the *ortho*-CF_3 substituent which substantially disfavors the WNA TS as well as the respective product. This observation resembles *Sabatier's* Principle which states that stabilization of a catalytic intermediate results in a larger barrier for the conversion to the product.[300] On the other hand, a methoxy substituent in *para* position appears to be particularly well suited for promoting the WNA. This is closely linked to the pyridine's increased basicity, which is strongest for *para*-OMe and favors the formation of a H−N bond in the TS and in the product. Interestingly, many of the substituents lead to somewhat smaller TS barriers compared to the non-modified pyridine, many in the range of 1 to 2 kcal mol^{-1} even when the modification makes the pyridine ligand less basic (e.g. *ortho*- and *meta*-OMe,[301] *para*-CF_3[302]). However, it can be seen that the basicity plays an important role in the stability of the WNA products (protonated pyridine) and that factors favoring the product decrease the TS energy. In a similar study Kang *et al.* calculated activation energies for a series of different ligand frameworks related to the famous Ru-bda catalyst by Sun and co–workers.[154] They report an inverse proportionality between the WNA barrier and the electrophilicity of the metal center. This is not the case for the modification of the Py5 system. However, pyridine is also not expected to affect the electrophilicity of the metal center since it is not directly coordinated to the metal center in the TS.

In Figure 5.11 the influence of the OMe, F and CF_3 substituents on the energies (of AR, TS and AP structures) and some properties of the free Ru^V=O-complex (Mulliken spin-densities on Ru and O and LUMO energies) is shown. Notably, the activation energies decrease for all three substituents going from *ortho*, *meta* to *para*. The strongest energetic effect of the substituents is observed for the AP structure, where a more basic pyridine leads to lower energies. The energy of the TS structure shows an intermediate dependence on the substituent position and basicity, reflecting the partially formed bond, while the ARs depend comparably little on the modification. Notice, how the spin-density on Ru is inversely correlated with the activation energy,

in contrast to the spin-density on oxygen. This intuitively reflects the change of electrophilicity – the less electrons there are on oxygen, the more prone it is to a nucleophilic attack. The correlation can also be seen with the LUMO energy, although the values show only small differences ($<3\,\mathrm{kcal\,mol^{-1}}$). The LUMO is localized entirely on the oxo–ligand, where it acts as the accepting orbital for the WNA.[175] However, these correlations are not necessarily transferable from substituent to substituent, i.e. the changes the spin-density do not quantitatively correlate with the changes in the activation energy etc. Nevertheless, a general energy ordering for the activation energies and product energies of *ortho* < *meta* < *para* can be established, which contrasts with the water-association reaction and showcases the different requirements of the reactions. As stated earlier solvation plays an important role in these reactions, and a different description might alter the obtained trends. Especially the energy from separated to ARs is, to a large extent, representing a solvation process. The importance of solvation has also recently been shown for RC mechanism of the WOCs designed by Sun and co–workers, where the oxo-ligand was found to be hydrophobic, thereby favoring the formation of the encounter complex prior to the O−O bond formation.[71,131,133]

Additionally, there is an intramolecular interaction between the electrophilic oxo and the pyridine fragment (denoted as L) in the $L-N^{4}-Ru^{V}-O-Cl$ structures, i.e. the reference states for AR, TS, and AP (see Table B.3). The shortest distance among the three substituents OMe, F, and CF_3 at the pyridine is found for the most basic substituent i.e. *para*-OMe. By adding the water molecules, this interaction is broken up. Its energy is therefore implicitly included in the energy difference between the separated and the ARs which may account for some part of the energy differences observed.

Substituent	Position	AR	TS	AP
unmodified		6.8	15.6	6.2
OMe	*ortho*	7.3	14.8	7.7
	meta	6.5	14.5	6.1
	para	6.2	11.7	3.2
F	*ortho*	7.2	17.1	15.5
	meta	6.8	14.7	10.4
	para	6.5	13.5	7.7
CF_3	*ortho*	9.2	20.9	17.1
	meta	7.3	16.3	11.6
	para	6.5	14.8	10.3
NO_2	*para*	6.8	15.3	12.0

Table 5.4: Energies of WNA AR, TS and AP structures in $\mathrm{kcal\,mol^{-1}}$. All energies are given with reference to the free complex and two water molecules. Energies were obtained at the B3LYP-D3/def2-TZVP/COSMO level of theory.

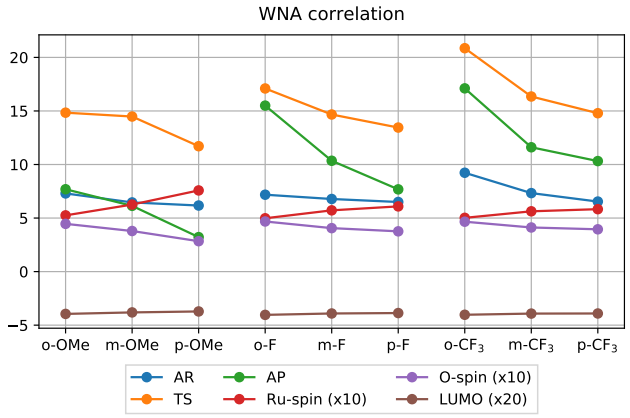


Figure 5.11: Correlation of quantities such as energies (AR, TS, AP; [kcal mol^{-1}]), spin-density on Ru and O (x10) and LUMO energy (x20 [eV]) of the free Ru^{V}=O species (OMe–N^{4}–Ru^{V}–O–Cl) with the position of the substituents on the axial pyridine.

Py5OMe Variant: para-OMe

As established in the previous section, a methoxy group at the axial pyridine of Py5OMe gives a WNA TS energy which is lower in energy than the unmodified variant. The water-association reaction on the Ru^{II} species is not substantially affected by this modification, as can bee seen from the thermochemical data given in Figure 5.12. Those values are compared to the unmodified ligand in order to assess the potential this ligand has for catalytic improvements. The water-association reaction on Ru^{III}is favored by approximately 2 kcal mol^{-1} with respect to the unsubstituted ligand, while the water-association reaction on Ru^{II} is slightly disfavored by -0.3 kcal mol^{-1}. While we have not generally observed a large dependence of the trends on the oxidation state, this example shows that a substituent can turn from having a slightly detrimental to a positive effect on the thermodynamics. Deprotonation reactions are slightly disfavored compared to the non-modified ligand. However, because of lower reduction potentials for III/II and IV/III couples, the PCET reactions require a similar amount of energy ($\pm$2 kcal mol^{-1}). Only the PCET from Ru^{IV}–OH to Ru^{V}=O is substantially larger due to the attenuated acidity of the complex. For a conceptional similar system it has been shown that changing the functional group coordinating to the metal center significantly alters the reduction potentials and thereby potentially the catalytic activity.[303] We further compared the thermodynamics of the whole catalytic cycle with an thermodynamically ideal catalyst as has been done in previous work.[67,131,133] Here, 'ideal' is meant in the sense that the Gibbs free energy associated with water oxidation is equally distributed among four PCETs, thereby minimizing the theoretical overpotential of the reaction. It becomes evident that the introduction of the *para*-OMe substituent barely affects the thermodynamics – as compared with the unmodified ligand (see Figure 5.13). This offers the opportunity to selectively alter the strength of the intramolecular base to lower the barrier for the WNA while at the same time keeping the favorable thermodynamics in place.

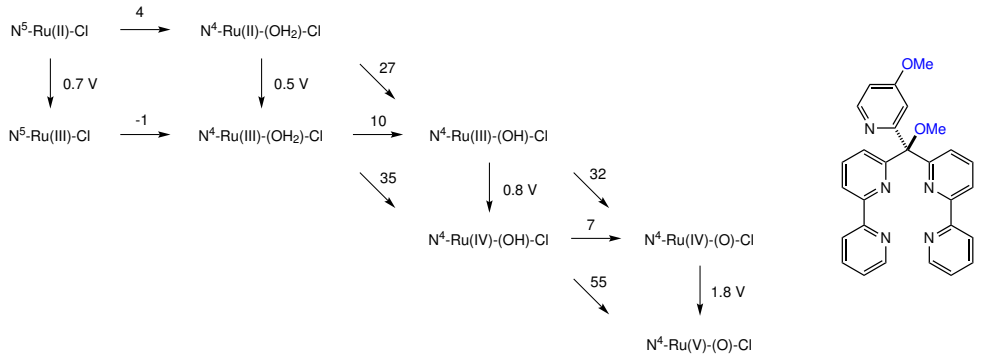


Figure 5.12: Energies of the reactions up to Ru^{V}=O for the *para*-OMe variant of Py5OMe. Energies are given in kcal mol^{-1}.

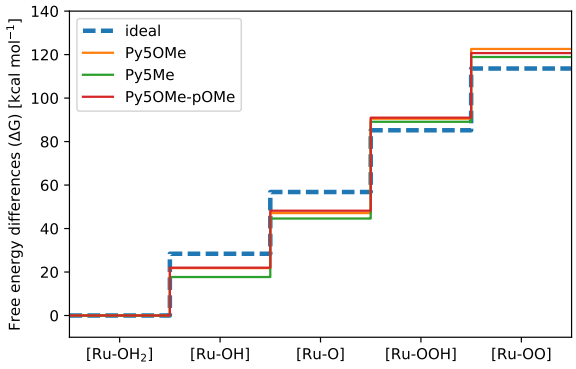


Figure 5.13: Thermodynamics of the water oxidation reaction catalyzed by Py5Me, Py5OMe, and Py5OMe variant *para*-OMe, compared to an 'ideal' catalyst.

5.3 Summary and Conclusions

Encouraged by our earlier study on Py5 derived Ru-based WOCs[173] we further explored the capabilities of the Py5 ligand framework. Clarifying the energetics of different coordination modes (i.e. isomers), we were able to solidify our previous observations based on Py5OMe and Py5Me, namely that they are in essence identical in terms of thermodynamics and kinetics (after the pyridine is replaced by a water molecule). The two ligands are primarily distinguishable by the sterical interaction of their substituents at the sp^3 carbon. This hypothesis was verified by studying a hypothetical Py5Et ligand. This catalyst fits perfectly in between the other two in terms of the thermodynamics of the pyridine-water exchange reaction as well as the kinetic barrier for the WNA, establishing the fact that small substituents at sp^3 carbon do not affect the water oxidation mechanism. The introduction of sterically demanding substituents of course would lead to significant destabilization of the ligand-metal interaction and potentially change in the favored coordination-mode. Focusing on the initial step prior to the water oxidation cycle, namely association of water to the catalyst, two possible routes were explored – the chlorido-water exchange and the pyridine-water exchange reaction. The first one is possibly part of a deactivation pathway and therefore unlikely part of the catalytic cycle. In agreement with experiments we find a correlation between the substitution of the sp^3

center and the activation barrier for the chlorido-water exchange. This is rather surprising since the substituent is too far apart to sterically interfere. Nonetheless, the Py5Me derived TSs appear to be much more strained compared to Py5OMe, in part due to the close proximity of the bipyridyl and the methyl. Dissociation scans of both the Ru–OH_2 and the Ru–N (pyridine) bonds suggest that modifications of the dangling pyridine may enhance the formation of the catalytically active Ru–OH_2 species. Performing *in silico* ligand design, we explored the influence of various functional groups and substitution patters on the Ru–OH_2 and Ru–N bond. Overall, the Ru–OH_2 bond is not significantly affected by the substitutions at the axial pyridine, while – depending on the position and nature of the substitution – the Ru–N bond can be altered efficiently. Those observations are illustrated by the correlation of the relative energies of R, AR, TS, and AP as well as observables from the electronic structure as a function of the substitution pattern. Introducing substituents in the *ortho* position imposed a sterical penalty which weakens the Ru–N bond. A decrease of the basicity caused by electron withdrawing groups further weakens the bond. However, in accordance with *Sabatier's* Principle we find that those substitutions disfavor the O–O bond formation by a WNA. On the other hand, *para* substituents, which only had a minor influence on the pyridine dissociation, lower the barrier for a WNA. At the same time, the thermodynamics remains virtually unchanged. Those findings suggest that modifying the basicity of the dangling pyridine can be done without disfavoring the other steps of the catalytic cycle. The Py5 system offers even more possibilities for modifications, in particular on the bipy fragments which are directly coordinated to the metal center. Alterations thereof are expected to have more pronounced influence on the thermodynamics of the catalytic cycle – however, the exploration of those modifications was beyond the scope of this work.

In this study we made a first step towards *in silico* design of Py5-ligands for water oxidation. A prerequisite for this is an in depth understanding of the catalytic cycle as well potential side reactions (isomerizations, ligand dissociations, etc.). It would be interesting to see whether changing the basicity of intramolecular bases would also primarily affect the WNA barrier in other systems and therefore become a general design concept.

Chapter 6

Determination of pKa Values via ab initio Molecular Dynamics and its Application to Transition Metal-Based Water Oxidation Catalysts

This chapter is a shorted version of: Schilling, M.; Luber S.; *Inorganics* **2019** *7*, 73.[98] If not stated otherwise all figures of this chapter were published in the reference mentioned beforehand.

A prerequisite for mechanistic studies is the knowledge of the chemical speciation of the transition metal complex under catalytic conditions. When working under aqueous conditions, special attention has to be given to functional groups which might undergo protonation / deprotonation reactions. Unfortunately it is often difficult to determine physical properties such as acidity constants (pK_a) of catalytic intermediates due to their elusive nature. In this study we apply DFT-MD, to reliably determine pK_a values of transition metal complexes used as catalysts for water oxidation (a brief summary of this chapter is given in Section 1.4.4). The here discussed methods have been successfully applied to a variety of compounds up to now, among them are organic molecules,[90,178,184] amino acids and peptides[179,185] as well as aqua complexes of transition metals.[182]

While all the studies mentioned above are based on the same protocol, namely the *Bluemoon* methodology, there are significant differences when it comes to the post-processing. In the following we will shortly introduce the general protocol, as well as the different flavors of post-processing. Then we compare those approaches for a set of benchmarking molecules, as well as for the system of interest - a ruthenium-based WOC $[Ru^{II}Py_5OMe(OH_2)]^{2+}$.[173] Said system and in particular the thermodynamics and kinetics of the water oxidation process were already investigated in-depth by Luber and co-workers employing DFT simulations (see Chapter 5.[173,174] Those studies gave important insights with respect to the water oxidation mechanism and further helped to come up with some design guidelines on how to further improve those catalysts. However, the limitations imposed by the implicit solvation led to the desire for a more sophisticated description of the solvation shell. In this context, also this study serves as a mini-review and benchmark study to validate the methodology for pK_a determination to be applied to the same or similar systems in the future.

6.1 Methodology

The calculation of pK_a values is a common task in computational chemistry - it is therefore not surprising that there are many protocols available for various levels of theory.[304] The most common approach used for small molecules is based on the calculation of the difference in free energy between the protonated and the deprotonated species using a thermodynamic cycle scheme, whereby the interaction between the solvent and solute is approximated by an electrostatic potential by means of an implicit solvation model.[61,62] Alternatively, the electrostatic contribution of solvation can be obtained in a two-step procedure. In the first step, the point charge distribution is generated by the restricted electrostatic potential (RESP) procedure.[305] Then the Poisson equation is solved in order to obtain the electrostatic energies of solvation.[306,307] In combination with sampling of the conformational space, the second approach is in particular useful for large systems such as proteins with multiple protonation sides.[308,309] Even though many protocols are able to reliably reproduce experimental pK_a values, their performance is still strongly system dependent. For example, large, flexible or highly charged species species stand in conflict with the underlying approximations of some of those protocols. There are numerous protocols and correction schemes depending on the system of interest which can be used to account for such shortcomings. A full review of them is beyond the scope of the current work, and we refer the interested reader to a number of selected articles.[60,75–77,79,310–313]

An obvious approach to improve the previously presented protocol is to describe the solvent at an atomistic level. However, if both the solute and the solvent are treated explicitly using AIMD not only the computational cost rises drastically but also the simulation protocols become more elaborated.[83,85,87,88,96,97,314]

Among those protocols is the so-called *Bluemoon* ensemble (see Section 2.4.1 for a brief introduction), where the free energy difference ΔF between the protonated and deprotonated state is calculated by a thermodynamic integration scheme. In this case the integrand is the average force ($f_{\xi'}$) acting on a system to impose the constraint (ξ), while the discrete values of the constraint (ξ') define the range of the integration:

$$\Delta F = -\int_{\xi_0}^{\xi_1} f_{\xi'} d\xi' \tag{6.1}$$

The latter is also referred to as potential of mean force (PMF). The average force $f_{\xi'}$ is derived from the Lagrange multiplier λ according to

$$f_{\xi'} = \frac{\langle Z^{-1/2}[\lambda - k_B T G]\rangle_{\xi'}}{\langle Z^{-1/2}\rangle_{\xi'}}. \tag{6.2}$$

where k_B is the Boltzmann constant, T the Temperature, and Z and G are correction factors associated with the transformation from generalized to Cartesian coordinates. In the case of distance constraint (d(A−H)) Equation 6.2 simplifies to $\langle\lambda\rangle_{\xi'} = f_{\xi'}$.

A detailed derivation of the *Bluemoon* methodology and in particular Equation 6.2 can be found in the corresponding original literature by Sprik and Ciccotti.[92–94]

The application of the *Bluemoon* methodology to determine pK_a values is in principle straight forward, however certain aspects deserve attention.

6.1.1 Choice of constraint

From a chemical point of view the two states of interest, i.e. the protonated and deprotonated species, are well defined. Either the proton is bound to the acidic functional group (A) or it is (infinitely) far away from it stabilized by an extensive hydrogen bonding network. Simulating the deprotonation state in a chemical sense is restricted due to the limitations with regard to the size of the simulation cell. There are sophisticated proton insertion schemes which avoid this problem.[83,85,87,88,90] However their high demand in terms of computational resources renders this approach unsuitable for complex systems.

A crucial choice within the *Bluemoon* methodology is the nature of the constraint that describes the two states. The most simple one is the A−H distance (d(A−H)). While it is often applicable, it suffers from some intrinsic problems. First, this constraint does not prevent the re-protonation of the acidic group. Secondly, upon deprotonation proton hopping might take place from the proton accepting molecule to other solvent molecules according to the *Grotthuss* mechanism. This is problematic since for large distances, in principle the intermolecular distance between the acidic group and a water molecule somewhere in solution is constraint. This is different from simulating a hydronium ion infinite separation from the acidic group. Besides that, the important question arises at which distance is a covalent O−H bond broken, or rather, which distances have to be sampled in order to reach the deprotonated state. In particular proton hopping makes this decision ambiguous since large values of d(A−H) may then only have the effect to constrain a solvent molecule in proximity to acidic group. Nevertheless there are several systems known for which the simple distance constrained turned out to work reasonably well.[95,178,179]

Those issues might be circumvented by not only constraining a single A−H bond but the CN of all the protons to the acidic group.[177] The CN is commonly represented as the sum over sigmoid functions such as a *Fermi-Dirac* distribution, and has been successfully applied to several systems.[90,96,97,183,184]

While a CN prevents the reprotonation of the acidic group it still does not prevent proton hopping. De Meyer *et al.* resolved this problem by constraining the difference in the CN of all the protons (index i, total number of protons: N) to the acidic group (index j) and a selected solvent molecule (index k) in its proximity:

$$CN(r_{ij}, r_{ik}) = \frac{\sum_i^N \left(1 - (\frac{r_{ij}}{r_0})^n\right)}{\sum_i^N \left(1 - (\frac{r_{ij}}{r_0})^m\right)} - \frac{\sum_i^N \left(1 - (\frac{r_{ik}}{r_0})^n\right)}{\sum_i^N \left(1 - (\frac{r_{ik}}{r_0})^m\right)} \tag{6.3}$$

where r_{ij} is the length of the vector $\overrightarrow{r_{ij}}$, describing the distance of proton i to the acidic group j, analogously r_{ik} is the distance of proton i to the solvent molecule k. The inflection point is defined by r_0, and exponential factors n and m define the overall shape of the switch function. This constraint in principle guaranties a smooth transition from the protonated acid to a solvent molecule without further proton hopping.[90] However said constraint introduces an additional empirical parameter, namely the choice of the proton accepting solvent molecule. Nonetheless, similar pK_a values were obtained when the accepting molecule was either part of the first or second solvation shell.[90]

6.1.2 Estimation of pK_a values from the free energy differences

Independent of the constraint used for the simulation, the free energy differences obtained from the *Bluemoon* ensemble might be interpreted as equilibrium constants from which pK_a values can be determined. There are several proposed ways, all of which have been shown to be able to reproduce experimental results reasonably well. However to the best of our knowledge those approaches have not be directly compared for the very same system. In the following we will present the different techniques, highlight their requirements in terms of simulations protocols and the necessary empiric parameters for the evaluation. For the detailed derivation of those methods, we refer the reader to the indicated references.

Absolute pK_a

The most straight forward approach is to use the relation between the equilibrium constant pK_a and the difference in free energy (ΔF):

$$pK_a = \frac{\beta \Delta F}{\ln(10)} \tag{6.4}$$

with $\beta = 1/k_B T$.

In order to reproduce experimental pK_a values one has to assure that the difference in free energy sufficiently describes the two states. This boils down to the nature of the constraint as well as to the range of sampling the corresponding phase space. In any case a key requirement is the convergence of $\Delta F(\xi)$ to constant value towards the limits of the sampled range (ξ_{min} to ξ_{max}), i.e. $\lim_{\xi \to \xi_{max}} \Delta F(\xi)$ = constant. Further, in order to obtain meaningful results from the thermodynamic integration the free energy at the bound state $F(\xi_0)$ has to be set to unity. This method has been successfully applied to several systems using either the d(A–H)−d(O–H), the CN(A–H) or CN(A–H)−CN(O–H) (see Equation 6.3) constraints.[90,178,180]

Whether the calculated free energy difference ΔF can directly be related to the equilibrium constant is disputed in literature.[97,182,185] Authors who disagree with the previously discussed method commonly refer to *Chandler's* derivations of the equilibrium constant based on a classical statistical mechanical description. The basic principle described there is the relation of the free energy difference ΔF and the radial distribution function (RDF) according to the reversible work theorem.[181] The RDF itself might be interpreted as the probability to find a proton within a certain radius of the acidic group. The probability distribution is related to the (inverse) acidity constant according to Davis *et al.*:[97]

$$K_a^{-1} = c_0 \int_0^\infty exp[-\beta \Delta F(r)] 4\pi r^2 \, dr \tag{6.5}$$

where c_0 is the standard concentration. Note the difference in free energy $\Delta F(r)$ is a function of radius of a sphere around the acidic group and not of the constraint ξ. Those two are only equal in case of the distance constraint d(A–H). In principle $\Delta F(r)$ has to be known for infinite separation, in practice however only a finite separation $R_{max} \leq L/2$ (L is the length of the cubic cell simulation) is accessible. Since $\Delta F(r)$ asymptotically approaches a constant value, one often defines $R_c \leq R_{max}$,

where R_c is the radius which distinguishes A−H from A^- + H^+, i.e. the distance at which the covalent bond is broken.[97] The limitations with respect to simulation cell results in an uncertainty in the pK_a value which Davies *et al.* quantified as $\Delta F(R_{max})/(2.3k_BT)$.[97]

Based on the RDF there are two common approaches to derive pK_a values, both of which require an additional set of simulations in order to reduce the potential errors describe above.

Relative pK_a

The approach presented by Ivanov *et al.* - in the later referred to as 'relative pK_a' - takes advantage of error-cancellation upon reporting the pK_a value relative to a pK_a value of a reference system (REF):

$$\frac{K_a^{HA}}{K_a^{REF}} = \frac{\int_0^{R_c} exp[-\beta\Delta F_{REF}(r)]r^2\,\mathrm{d}r}{\int_0^{R_c} exp[-\beta\Delta F_{HA}(r)]r^2\,\mathrm{d}r}. \tag{6.6}$$

Compared to Equation 6.5 the integration is from 0 to R_c ($R_c \leq R_{max}$) in order to account for the fact that $\lim_{r\to R_{max}} \mathrm{p}K_a(r)$ asymptotically approaches p$K_a(R_{max})$, i.e p$K_a(r)$ quickly becomes a constant value.[185] Further, the lower bound of the integral can be approximated by the value slightly smaller than the average A−H bond length, since the contributions of large values of $-\Delta F(r)$ in the exponent are negligible.[185] The reference system (see Equation 6.6) has to be calculated within the same computational framework, which in fact doubles the cost of the approach compared to the previously described method. However, the reference system might be used for the determination of pK_a values of multiple species. Further in order to equate Equation 6.6, $\Delta F(R_{max})$ are set to unity for both the acid and the reference system.[179,185]

This method has been successfully applied to several isomers of histidine as well as a histidine-tryptophan dimer employing a simple d(A−H) constraint.[179,185]

Probabilistic pK_a

The second approach was introduced by Davies *et al.*, later referred to as 'probabilistic pK_a'.[97] The main idea is to define the acid dissociation constant K_a by the probabilities to find protons within a certain volume of the acidic group. Then the normalized probability to find a proton within a cutoff radius R_c is

$$P(R_c) = \frac{\int_0^{R_c} exp[-\beta\Delta F(r)]r^2\,\mathrm{d}r}{\int_0^{R_{max}} exp[-\beta\Delta F(r)]r^2\,\mathrm{d}r}. \tag{6.7}$$

where R_c is the cutoff radius defining the protonated and the deprotonated state and R_{max} the limit for infinite separation.

In the limit for weak acids the K_a value then becomes

$$K_a(R_c) = \frac{(1-P(R_c))^2}{P(R_c)}\frac{N}{c_0 V} \tag{6.8}$$

where c_0 is the standard concentration, V the volume of the simulation cell and N the number of acidic sides - here N = 1. The latter is converted to mol by dividing it by the Avogadro constant.

Here again the choice of the cutoff radius R_c is crucial. However unlike the relative pK_a scheme, there is no obvious asymptotic behavior within the range of the deprotonation. This led Davies *et al.* to calculate the p$K_w$of liquid water using the following relation:

$$K_w(R_c) = \left((1 - P(R_c))^2 \frac{N_w}{c_0 V} \right) \tag{6.9}$$

where N_w is the number of water molecules in the simulation cell. R_c is the radius at which $pK_w(r) = 14$ is true.[97] The relation between Equation 6.9 and Equation 6.8 is that the activity of the undissociated reactant is set to unity.[97]

As for the previous methods, the current one has been successfully applied to several systems employing either a simple d(A – H) constraint or more commonly a CN constraint.[96,97,182–184]

By introducing the three approaches named 'absolute', 'relative' and 'probabilistic' pK_a we have laid the foundation for the following study.

6.2 Computational Settings

All AIMD simulations were performed at the BLYP-D3/DZVP-MOLOPT-SR-GTH level of theory (see Section 2.5.2) employing the CP2K program package.[241] In order to enlarge the time step the mass of all hydrogen atoms was set to 2 a.m.u. in accordance with literature.[90] The influence of which is discussed in Section 6.3. Further a cutoff of 800 Ry for the auxiliary plane wave basis set was used.

In order to simulate liquid water, we used a cubic simulation cell with a side-length of 15.6404 Å containing 128 water molecules. This box size corresponds to liquid water at 1 bar and 300 K using the TIP5P force field.[315] The same simulation cell was used for solvated molecules, by removing several water molecules we assured that the pressure remained approximately the same as the one of clean liquid water. For some model systems a larger simulation cell with side-length 19.7340 Å containing 256 water molecules was employed.

The simulations were performed in the NVT ensemble with a time-step of 0.5 fs. The temperature was kept constant at 320 K by a Nosé-Hoover chain thermostat.[210,211] The slightly elevated temperature is required in order to avoid the glassy behavior of BLYP water.[316]

The general settings mentioned above closely resemble the protocols that have been employed previously by other groups in order to determine pK_a values.[90,179,185]

Each model system was equilibrated in the protonated state for 5-10 ps without any constraint. Starting from those simulations, constrained AIMD runs were performed each for an additional 15-20 ps (30000-40000 steps (see Tables C.11 and C.12). The first 5 ps (10000 steps) of each individual run were neglected in order to give the system time to equilibrate e.g. adopt to the imposed constraint. The convergence of each model system with respect to simulation time is given in the Appendix (Tables C.3, C.4, and C.5).

Model Systems

Simulations were conducted for the systems given in Table 6.1. pK_a values were not only calculated for transition metal complexes, but also for two small organic molecules with pK_a values in the same range as the molecules of interest. Thereby phenol serves as an internal standard which allows for direct comparison with pK_a values obtained by the same methodology.[90]

Table 6.1: Model systems used in this study. N_w stands for the number of water molecules in the simulation cell. The calculated pK_a value is only given if the simulations found in the literature were obtained employing the *Bluemoon* methodology. The Ru complex bears a Py5 ligand that is composed of two bipy fragments linked to a fifth py via an sp^3 carbon. The fourth fragment connected to the latter is either a methyl (Me) or methoxy (OMe) group resulting in Py5Me or Py5OMe (see Chapter 5 for more details).[172–174] Experimental pK_a values denoted with "*" were only available for the Py5Me ligand framework (see Figure 1.15 for a graphical representation of the catalysts).

Molecule	N_w	side-length [Å]	pK_a (exp.)	pK_a (calc.)
H_2O	128	15.6	14.0	-
H_2O	256	19.7	14.0	-
HCOOH	126	15.6	3.8[317]	-
PhOH	123	15.6	10.0[317]	9.7[90]
$[Ru^{II}Py_5Me(H_2O)]^{2+}$	112	15.6	∼11[173]	-
$[Ru^{II}Py_5OMe(H_2O)]^{2+}$	112	15.6	∼11[173]*	-
$[Ru^{II}Py_5OMe(H_2O)]^{2+}$	234	19.7	∼11[173]*	-
$[Ru^{III}Py_5OMe(H_2O)]^{3+}$	112	15.6	∼3[173]*	-
$[Ru^{III}Py_5OMe(H_2O)]^{3+}$	234	19.7	∼3[173]*	-

In our previous study we have investigated the water oxidation mechanism of $[Ru^{II}Py_5Me(H_2O)]^{2+}$ and $[Ru^{II}Py_5OMe(H_2O)]^{2+}$ listed in Table 6.1, from a kinetic and thermodynamic point of view employing state of the art DFT simulations (see Chapter 5).[173,174] We found that both ligand frameworks Py5OMe and Py5Me were virtually identical in terms of their thermodynamics and kinetics of the water-oxidation reaction, which is a not too surprising result as the replacement of a methyl-group by a methoxy-group at an sp^3 carbon is not expected to remarkably alter electronics or sterics at the metal center. However, experimentally the catalytic activity between the two ligands was found to be rather different, which was attributed to rapid halide substitution of the Py5Me ligand compared to the Py5OMe ligand framework. This leads to a deactivation of the catalyst.[173] As mechanistic studies for the more active catalyst Py5OMe are still underway, we decided to choose it as a model system even though no experimental pK_a values are currently available. Based on our previous study we would not expect the thermodynamics of the two ligands and related pK_a values to differ significantly (see Chapter 5).[13,173] Therefore comparing the calculated pK_a value of the Py5OMe and Py5Me systems serves as a further validation of the method.

Error Analysis

The standard deviation of the average forces used to calculated the free energy difference according to Equation 6.1 is calculated by block averaging methods.[318] An upper limit for the standard deviation (σ) of the pK_a is obtained by calculating the free energy ΔF of both, the average force ($\langle\lambda\rangle$) and the average force plus its standard deviation i.e. ($\langle\lambda\rangle + \sigma_\lambda$).[95]

6.3 Results and Discussion

6.3.1 Convergence of the AIMD Simulations

For our comparison of the post-processing methods we ran constrained AIMDs for the systems described in Section 6.2. As a constraint we chose the distance of the acidic proton from the acidic group (d(A–H)), which we scanned in steps of 0.1 Å over a range from 0.9 Å to 1.6 Å. The constraint was primarily chosen for its simplicity, which avoids additional parameters, i.e. the explicit definition of the proton accepting molecule. Further, Equation 6.1 holds only for the distance constraint, otherwise correction terms are necessary (see Equation 6.2).[92] For all the systems, proton hopping is observed for d(A–H) distances larger than 1.5 Å, which made the scanning of distances larger than 1.6 Å obsolete. In particular since large values of d(A–H) do not necessarily describe the distance between the acidic group and the initially formed hydronium, since the latter potentially loses its one of its proton to other solvent molecules.

An exemplary PMF profile of PhOH is shown in Figure 6.1. Error bars on the average forces were obtained by block averaging methods. The absolute value of those standard deviations is in the range of 0.5 kcal mol^{-1} to 1.3 kcal mol^{-1}, which corresponds to 0.3 to 0.9 pK_a units. The PMF profiles of all the other model systems can be found in the Appendix: Figures C.1 to C.8.

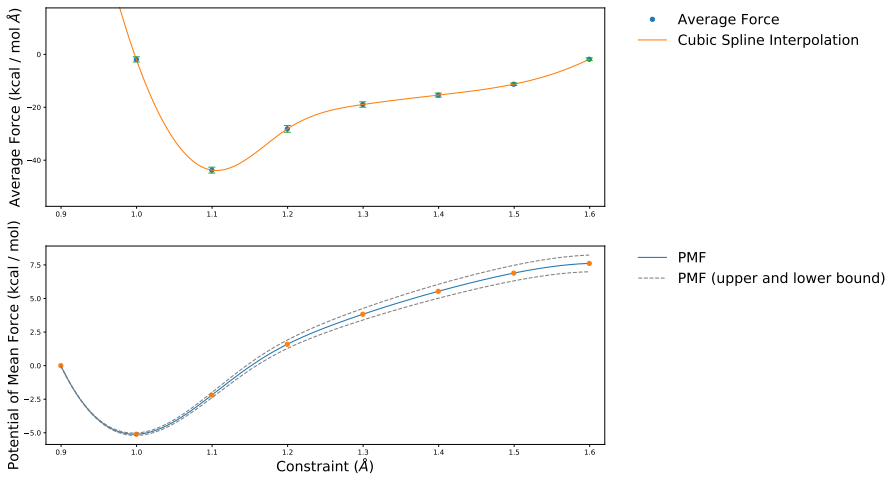


Figure 6.1: Top: Average force ($\langle\lambda\rangle$) acting on the constraint for the PhOH model system, constraining the A−H bond. Bottom: Potential of mean force, i.e. free energy obtained by integrating the average forces.

The convergence of the force acting on the constraint can be seen in Figure 6.2. Each of the traces represents a single point in the top part of Figure 6.1. Towards larger values of the constraint the fluctuations start to oscillate around zero, highlighting the fact that the A−H bond has been broken.

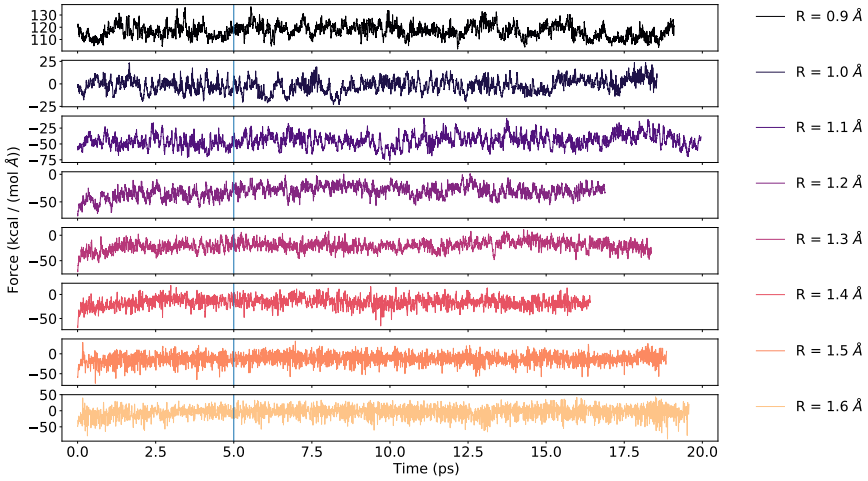


Figure 6.2: Force (λ) acting on the constraint - here for PhOH at a constrained d(A−H). The vertical line at 5.0 ps marks the equilibration time in relation to the production run.

The autocorrelation function of the force acting on the constraint is used to deter-

mine the optimal block size for the average which is in the range of 0.5 to 1 ps (see Figure 6.3). The force appears to be heavily correlated for the constraints $\leq$ 1.3 Å which corresponds to the region where the A−H bond is broken (see Figure 6.2). Ivanov *et. al* further analyzed those autocorrelation functions in order to elucidate the bond breaking process.[185] Since the thermostat modulates the autocorrelation function, simulations in the NVE ensemble would be required which are beyond the scope of the current work.

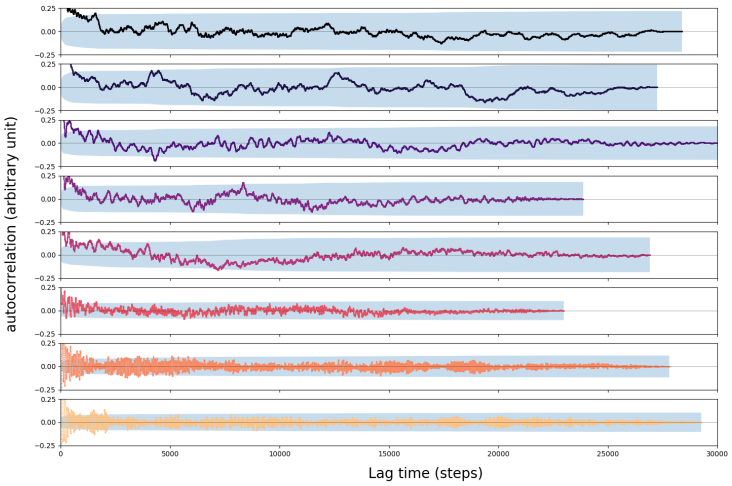


Figure 6.3: Autocorrelation function of the force acting on the constraint. From top to bottom the constraint increases by 0.1 Å, starting from 0.9 Å.

Reference System

In order to calculate the pK_a value according to the 'relative' or 'probabilistic' method, the simulation of water is required. We determined a cutoff radius R_c of 1.24 Å according to Equation 6.9 (see Figure 6.4). We determined the same value for the 128 and 256 water molecule simulation cell (see Figure C.16), which is in good agreement with 1.22 Åand 1.28 Å , reported in literature.[96,97,182] As already reported by others, strong bases such as OH^- tend to be reprotonated by the solvent. This is also the case in our simulations. Starting from a d(O−H) of 1.5 Å, we were able to observe the reprotonation of the OH^- moiety (see Figure 6.5). As discussed earlier, choosing a different collective variable could circumvent this issue.

In the following we will present the pK_a values calculated based on the protocol described beforehand.

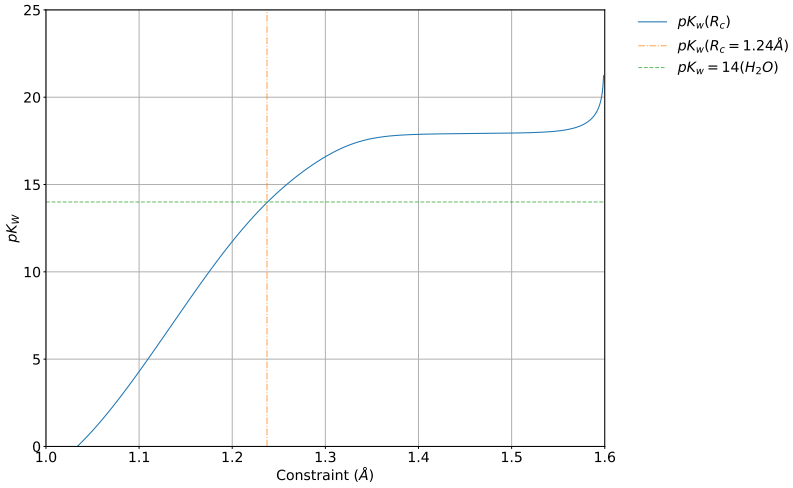


Figure 6.4: Determination of R_c from simulation of water, by fitting to the experimental value. The simulations were carried out in a cubic box with a side length of 15.6406 Å.

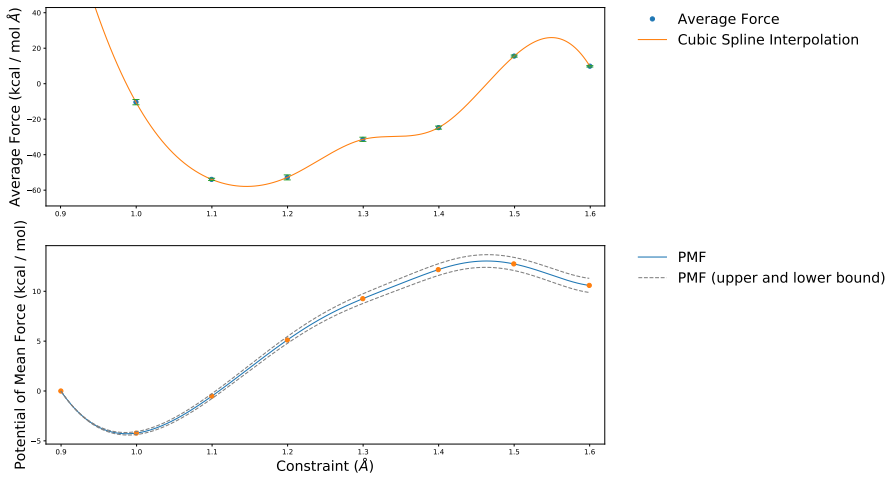


Figure 6.5: Autodissociation of H_2O - Top: Average force acting on the constraint. Bottom: Free energy obtained by integrating the average forces. Note the drop in free energy at 1.5 Å is caused by the reprotonation of OH^-.

6.3.2 Overview of calculated pK_a values

In Table 6.2 the pK_a values obtained with the three simulation protocols are given. In general we find that the 'absolute' protocol (see Equation 6.5) underestimates the pK_a values compared to the experiment by about 1-2 pK_a units. pK_a values obtained by the relative protocol on the other hand are an overestimation by about 2 pK_a units. The best agreement is achieved with the probabilistic protocol where

in particular pK_a values in the range of 10-11 are accurately reproduced. Hereby $[Ru^{III}Py_5OMe(H_2O)]^{3+}$ is somewhat of an outliner with the largest difference to the experiment by 2 pK_a units. As the only difference between $[Ru^{II}Py_5OMe(H_2O)]^{2+}$ and $[Ru^{III}Py_5OMe(H_2O)]^{3+}$ is the charge of the system, we reasoned that the simulation cell might be too small for such highly charged species. This was verified by a set of simulations in a bigger simulation cell (see Table 6.3).

Table 6.2: All results presented here are calculated at 320 K, for a cubic box with side length of 15.6406 Å and a cutoff R_C = 1.24 Å, over a trajectory of about 20 ps (where the first 5 ps were not included in the evaluation). The standard deviation is calculated using the block average method with a block size of 1 ps; [a] absolute, [b] relative, [c] probabilistic protocols.

Molecule	pK_a (exp.)	pK_a[a]	pK_a[b]	pK_a[c]
H_2O	14.0	-	-	
HCOOH	3.8[317]	2.7 ± 0.5	6.7 ± 0.5	4.2 ± 0.6
PhOH	10.0[317]	8.7 ± 0.3	12.5 ± 0.5	10.7 ± 0.4
$[Ru^{II}Py_5Me(H_2O)]^{2+}$	∼11[172]	9.8 ± 0.4	13.7 ± 0.3	11.2 ± 0.4
$[Ru^{II}Py_5OMe(H_2O)]^{2+}$	∼11[172]	9.3 ± 0.4	13.3 ± 0.6	11.1 ± 0.4
$[Ru^{III}Py_5OMe(H_2O)]^{3+}$	∼2.5[172]	3.1 ± 0.4	7.1 ± 0.4	4.5 ± 0.5

The larger simulation cell slightly improves the agreement between the calculated pK_a values using the relative protocol and the experiment. However, there is no systematic improvement of the pK_a values with the system size, independent of the overall charge. Due to the shorter trajectories as compared to the smaller system (see Tables C.11 and C.12), we could not obtained the same level of convergence, which can be seen from the standard deviation.

Table 6.3: All results presented here are calculated at 320 K, for a cubic box with a side length of 19.7340 Å and a cutoff R_C = 1.24 Å over a trajectory of 10-15 ps (where the first 5 ps were not included in the evaluation). The standard deviation is calculated using the block average method with a block size of 1 ps; [a] absolute, [b] relative, [c] probability protocols.

Molecule	pK_a (exp.)	pK_a[a]	pK_a[b]	pK_a[c]
H_2O	14.0	-	-	
$[Ru^{II}Py_5OMe(H_2O)]^{2+}$	∼11[172]	10.1 ± 0.5	11.5 ± 0.5	12.7 ± 0.7
$[Ru^{III}Py_5OMe(H_2O)]^{3+}$	∼2.5[172]	3.1 ± 0.3	4.6 ± 0.3	4.3 ± 0.4

6.3.3 Deuterated Solvent

The use of deuterated solute and solvent i.e. setting the mass of hydrogen atoms to 2 a.m.u has a profound impact on calculated pK_a values. In principle all values presented in Tables 6.2 and 6.3 are p$K_a{}^D$ values, i.e. pK_a values in D_2O instead of p$K_a{}^H$ values, i.e. pK_a values in H_2O. Based on experiments a correlation between p$K_a{}^D$ and p$K_a{}^H$ was reported already decades ago[319]. At first, the correlation was suspected to be linear only for pK_a values ¿ 7 and constant for more acidic species.[319]

Later, Delgado *et al.* experimentally determined a linear relation between $pK_a{}^D$ and $pK_a{}^H$ over the whole range of pK_a values:[320]

$$pK_a^D = 1.044 pK_a^H + 0.32 \tag{6.10}$$

In a more recent study a detailed derivation of the linear relation between $pK_a{}^D$ and $pK_a{}^H$ has been presented by Krężel and Bal:[321]

$$pK_a^H = 0.929 pK_a^{H^*} + 0.41 \tag{6.11}$$

where $pK_a{}^{H^*}$ is the pK_a value determined in a D_2O solution by a pH-meter which was calibrated by H_2O. Conversion to $pK_a{}^D$ is achieved by adding the empirically determined constant of 0.4 to $pK_a{}^{H^*}$:[321]

$$pK_a^D = pK_a^{H^*} + 0.4 \tag{6.12}$$

Combining Equation 6.12 and Equation 6.11 and solving for $pK_a{}^D$ results in:

$$pK_a^D = 1.076 pK_a^H - 0.041 \tag{6.13}$$

which is of the same mathematical form as Equation 6.10. For the probabilistic method, further the cut-off radius R_c has to be redetermined since the $pK_w{}^D$ of D_2O is 14.951 (25° C).[322] The obtained values for R_c are 1.26 Åand 1.25 Å for the two simulation cells (see Figures C.17 and C.18). The $pK_a{}^D$ values obtained with the cut-off R_c determined for D_2O were converted to $pK_a{}^H$ values (see Tables 6.4 and 6.5).

Table 6.4: All results presented here are calculated at 320 K, for a cubic box with a side length of 15.6406 Å, over a trajectory of about 20 ps (where the first 5 ps were not included in the evaluation). The pK_a values were calculated using the probabilistic method for a R_c value of 1.26 Å determined for D_2O. The $pK_a{}^H$ [a] was obtained referencing the calculations to H_2O i.e a R_c of 1.24 Å. The $pK_a{}^D$ values were converted to $pK_a{}^H$ values according to Equation 6.10 [b] and Equation 6.13, respectively [c].

Molecule	pK_a (exp.)	$pK_a{}^H$[a]	$pK_a{}^D$	$pK_a{}^H$[b]	$pK_a{}^H$[c]
H_2O	14.0	-	-	-	-
HCOOH	3.8[317]	4.2	4.3	3.8	4.0
PhOH	10.0[317]	10.7	11.2	10.4	10.5
$[Ru^{II}Py_5Me(H_2O)]^{2+}$	∼11[172]	11.2	12.0	11.2	11.2
$[Ru^{II}Py_5OMe(H_2O)]^{2+}$	∼11[172]	11.1	11.8	11.0	11.0
$[Ru^{III}Py_5OMe(H_2O)]^{3+}$	∼2.5[172]	4.5	4.6	4.1	4.3

The probabilistic $pK_a{}^H$ values obtained by applying the conversion schemes discussed above are very similar to pK_a values determined by referencing our calculations to H_2O instead of D_2O. The difference when converting the $pK_a{}^D$ values either according to Equation 6.10 or Equation 6.13 is negligible.

Table 6.5: All results presented here are calculated at 320 K, for a cubic box with a side length of 19.7340 Å, over a trajectory of about 20 ps (where the first 5 ps were not included in the evaluation). The pK_a values were calculated using the probabilistic method for a R_c value of 1.25 Å determined for D_2O. The p$K_a{}^H$ [a] was obtained referencing the calculations to H_2O i.e a R_c of 1.24 Å. The p$K_a{}^D$ values were converted to p$K_a{}^H$ values according to Equation 6.10 [b] and Equation 6.13 respectively [c].

Molecule	pK_a (exp.)	p$K_a{}^H$[a]	p$K_a{}^D$	p$K_a{}^H$[b]	p$K_a{}^H$[c]
H_2O	14.0	-	-	-	-
$[Ru^{II}Py_5OMe(H_2O)]^{2+}$	∼11[172]	12.7	13.1	12.2	12.2
$[Ru^{III}Py_5OMe(H_2O)]^{3+}$	∼2.5[172]	4.3	4.4	3.9	4.1

The p$K_a{}^D$ or rather the p$K_a{}^H$ values were calculated either by the absolute or relative method see Tables C.7 to C.10. For the relative protocol, the overestimation p$K_a{}^H$ values is reduced by about 1-2 units. However, no systematic improvement is achieved as in particular the p$K_a{}^H$ values for the acidic compounds were still significantly overestimated. The agreement between the experimental and calculated p$K_a{}^H$ values using the absolute protocol decreased by 0.1 to 0.6 units upon converting p$K_a{}^D$ to p$K_a{}^H$ values.

Taking into account the multitude of empirical factors required to convert the p$K_a{}^D$ values led to the conclusion that the referencing pK_a values to H_2O is acceptable. This is in particular true for the probabilistic method. The validity of this conclusion could in principle be checked by repeating all the simulations with H_2O instead of D_2O. With a large enough test set, it would also be possible to adjust equation of the linear relation to the employed methodology. However this is beyond the scope of the current work.

In the following we are going to highlight the dependence of the three methods on the cut-off radius (R_c).

6.3.4 Absolute and Probabilistic pK_a - dependence on R_c

In Figure 6.6 both the absolute as well as the probabilistic pK_a values are shown as a function of the constraint (see Figures C.9 to C.15 for plots for the other model systems). The choice of R_c is obviously crucial (see Figure 6.6). In order to illustrate the influence of R_c, pK_a values were also calculated with R_c values reported in literature (see Tables C.1 and C.2). Some systems are remarkably independent of R_c while for others they spread over a range of almost 3 pK_a units. This suggests that it is crucial to determine R_c with exactly the same settings as the systems of interest, rather than to rely on an previously published value.

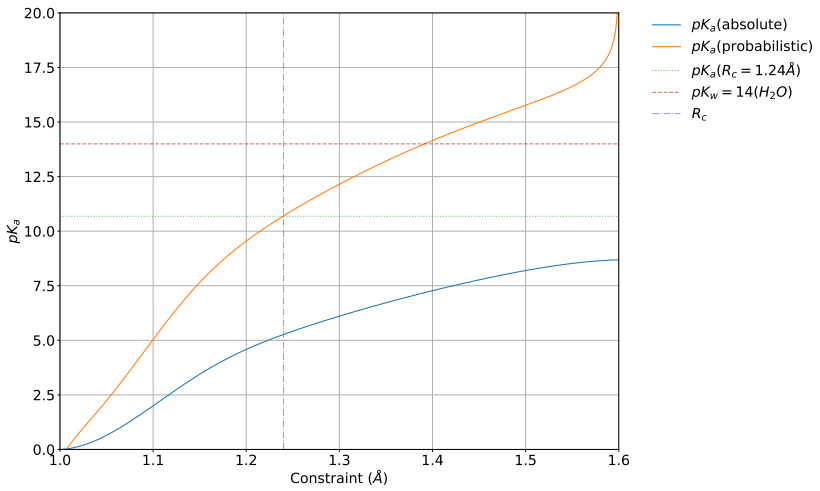


Figure 6.6: pK_a values calculated for PhOH using the absolute and probabilistic method. The latter is strongly dependent on the choice of R_c, even a change by 0.01 Å might result in change of up to 0.5 pK_a unit.

6.3.5 Relative pK_a

As for the probabilistic protocol, the dependency of the relative pK_a on R_c was investigated. As expected, the pK_a values calculated by Equation 6.6 show an asymptotic behavior towards larger values of R_c (see Figure 6.7).[185] This suggests that large values of R_c might be neglected in the evaluation. Indeed the results can be tuned i.e. lowered by 1 pK_a unit, if only constraints between 0.9 and 1.4 Å are considered (see Table C.6). Nonetheless the results for the acidic compounds are still overestimated with 5.8 and 6.2 for HCOOH and $[Ru^{III}Py_5OMe(H_2O)]^{3+}$ respectively. Inconsistency among different systems have also been observed by Kiliç *et al.* when investigating lumiflavins.

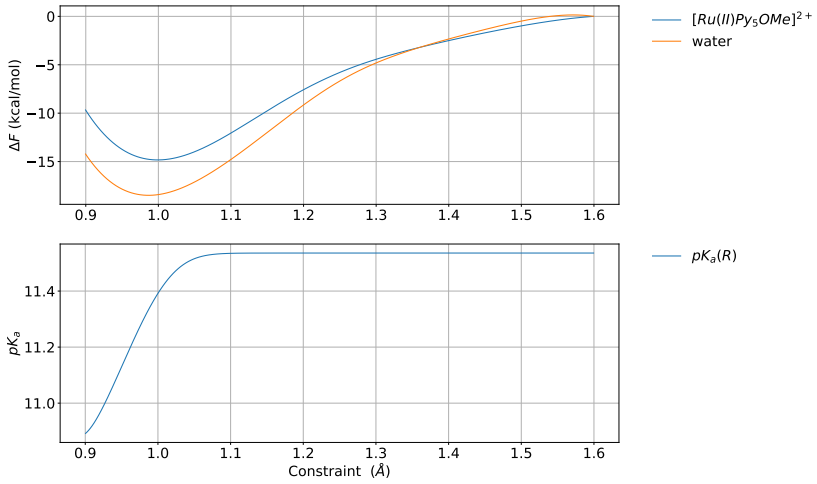


Figure 6.7: Relative pK_a as a function of d(A−H) according to Equation 6.6, the value asymptotically approaches a constant value.

6.4 Summary and Conclusion

In summary we have investigated the Bluemoon methodology and its various post-processing methods in order to determine pK_a values previously only applied to small organic molecules and transition metal aqua complexes. Our simulation cells with 128 or 256 water molecules were considerably larger than the ones previously presented in literature. This highlights the robustness and applicability of said protocol to larger chemically interesting systems, in particular in the context of WOCs such as the here presented $RuPy_5Me$ and $RuPy_5OMe$ catalysts, for which we found, as expected, both qualitative and quantitative similar pK_a values independent of the applied post-processing method.

When comparing the three post-processing methods suggested in literature (1) absolute, (2) relative and (3) probabilistic pK_a (all referenced to H_2O), we find that method 1 and 3 are able to quantitatively reproduce experimental values with an accuracy of about 1 pK_a unit. In case of method 1 this is rather surprising as there is no guarantee that the free energy levels off within the scanned range of the constraint. We further want to stress out, that the choice of a simple distance constraint might be optimal for this protocol. Method 2 appears to at least qualitatively reproduce experimental pK_a values, however there seems to be no necessity to use said approach, as it requires exactly the same set of calculations as method 3 which preformed the best within our test set. This conclusion also holds if one assumes that $pK_a{}^D$ instead of $pK_a{}^H$ values were calculated. For the sake of consistency we have only applied a very simple constraint i.e. the A−H distance. While this choice was fine for our test-set, there might be cases where more complex constraints are necessary, in particular if the conjugated base is very strong.

The overall accuracy of pK_a values calculated by the Bluemoon methodology using a simple distance constraint lies between 1-3 pK_a units. However if the computational settings are tailored in order to reproduce relevant experimental reference values,

the accuracy might be increased. Overall it appears that both methods 2 and 3 slightly overestimate small pK_a values while large pK_a values are reproduced with high fidelity. This issue should be addressed in an extended benchmark study before one attempts to accurately predict low absolute pK_a values. For method 1 on the other hand most pK_a values were underestimated by 1-2 units.

Further we want to highlight the importance of a sufficiently large simulation cell when investigating highly charged systems such as WOCs with different oxidation states. While the influence on the calculated pK_a values was found to be minor, only a large enough simulation cell can guarantee that there are no spurious interactions between the solute and its mirror image in the neighboring simulation cells.

Chapter 7

Zooming in on the O−O Bond Formation - an ab initio Molecular Dynamics Study Applying Enhanced Sampling Techniques

This chapter is a shorted version of: Schilling, M.; Cunha R. A.; Luber S.; *submitted*

In a previous Chapter 5 we throughly studied the mechanism governing water oxidation catalyzed by {Ru(Py5OMe)}. There the intermediates as well as the corresponding TSs were simulated by means of DFT calculations, whereby the energetic contributions of the solute-solvent interaction were approximated by the COSMO.[59,205] Within said method directed solute-solvent interactions such as hydrogen bonding are not explicitly accounted for. This can have profound consequences on the obtained structures. For example, in the structure of the R, a Ru^{V}=O species, the lone pair of the pyridyl is aligned with the oxo-ligand, implying weak electrostatic interactions between them (see Figure 7.1. Another limitation of the method became evident while attempting to model the TS. Localization of them was unsuccessful unless at least one additional water molecule was included in the model system. The inclusion of a limited amount of explicit solvent molecules is a common strategy when it comes to modeling TSs and MEPs connecting two states.[35,67,131] The limitations of this approach have been investigated by Hodel *et al.* in the context of ligand exchange reactions on a cobalt based WOC, where the whole first solvation sell was treated explicitly and compared to metadynamics simulations.[68]

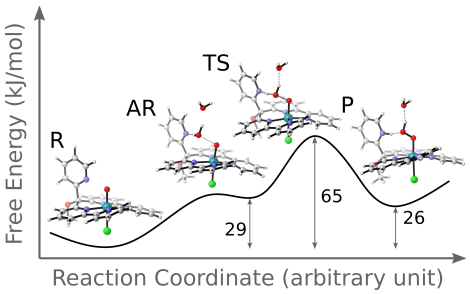


Figure 7.1: Visualization of the relevant intermediates of the O−O bond formation catalyzed by {Ru(Py5OMe)}. The structures are reproduced from geometrical information published by Gil-Sepulcre *et al.* and so are the relative free energies.[173]

The inclusion of an additional solvent molecule in the static simulation of the TS of the WNA of {Ru(Py5OMe)} had also deeply affected the structures of the AR and P, the other extrema of interest. The inability of the protic groups i.e. the water molecules or the newly formed hydroperoxo ligand, to engage in hydrogen bonding with surrounding solvent resulted in a overly stable hydrogen bonding network that is conserved among the three structures of the extrema (AR, TS, P). While this observation is expected within the applied method, there is no foundation to expect this to be true for the 'real' system. The shortcomings of this approach with respect to the description of the solvent dynamics and conformational diversity of the states of interest therefore need to be addressed.

In this Chapter we improve the understanding of the WNA on {Ru(Py5OMe)} by applying state of the art DFT-MD, usually called AIMD simulations (a brief summary of this chapter is given in Section 1.4.4). Since the O–O bond formation is an activated process, enhanced sampling methods such as the *Bluemoon* ensemble and MetaD were used. In the following we are going to discuss in-depth how to apply those methods to a complex chemical system in order to elucidate the role of the solvent arrangement, dynamics, and the intramolecular base in the O–O bond formation mechanism. Similar studies employing AIMD enhanced sampling are known in the literature for analogous reactions such as the formamide dehydrogenation, the trifluoromethylation of thiols and acetonitrile or the capture of CO_2 by frustrated lewis pairs.[84,323–325] However, studies in the context of water oxidation are rare. One of the earliest studies dates back to 2010 and was conducted by Vallés-Pardo *et al.* They used MetaD to simulate an initial reaction path of the O–O bond formation by a WNA which was then refined by *Bluemoon*.[69] An analogues approach was also used by Piccinin *et al.* which will be discussed later on in more detail.[70] In a recent study Govindarajan *et al.* used *Bluemoon* to model the a WNA. In their case, a solvent molecule acted as a proton relay between the nucleophile and a dangling carboxylate, a situation analogous for the WOCs studied in this work.[37] Employing empiric valence bond MD (EVB-MD)[73] Zhang *et al.* proposed for the same WOC as the one studied by Govindarajan, a nucleophilic attack of a hydroxide onto the carboxylate which then transfers one of its oxygen atoms to the metal-oxo species forming the O–O bond.[41] Oxygen-atom transfer reactions such as the one discussed before have recently been suggested as an alternative to the classical WNA.[39] To this extent Govindarajan *et al.* studied the formation of a of pyridyl-N-oxide i.e the transfer of the metal-oxo ligand to a dangling pyridyl (N_{py}–O) by constrained DFT-MD (*Bluemoon*).[40] This species is then expected to form the O–O bond formation. An analogue mechanism could also be imagined for the Py5-based WOC, but has not been considered so far in previous studies and is beyond the scope of the current study. In 2019, Shao *et al.* modeled the catalytic cycle of a Ru-based WOC. The ligand of which is covalently linked to an organic dye that upon excitation acts as an intramolecular oxidant.[326] They used constrained DFT-MDs to model the deprotonation reactions, which are expected to take place during the catalytic cycle, as well as the O–O bond formation by a WNA. The formation of the O–O bond together with the subsequent deprotonation of the nucleophile was found to be the energetically most demanding reaction of the catalytic cycle.[326] In all of those reactions the solvent molecules accepted the protons released from the catalyst, underlining the need for an explicit treatment of the solvent. All the works introduced before apply rather simple order parameters to describe the O–O bond formation, usually the O–O distance. In this work we show how that the use of simple order parameters

might leads to a qualitatively wrong picture of the underlying free energy surface. In addition, using a sophisticated amount of solvent molecules in the model system demonstrates the essential role of the solvent environment.

7.1 Methods

In this work the enhanced sampling methods the *Bluemoon* and MetaD were used to model the O−O bond formation. The *Bluemoon* method is briefly introduced in Section 2.4.1, for a detailed derivation see the corresponding original literature by Sprik and Ciccotti.[92–94] Analogously, the MetaD method is described in Section 2.4.2, for a detailed derivation we refer to the original literature and the corresponding review articles.[45–47,220,221]

7.1.1 Computational Settings

All AIMD simulations were performed employing the CP2K program package (revision 18461).[241] Ahe electronic structure is described at the PBE-D3/DZVP-MOLOPT-SR-GTH level of theory,[192,199,327] for more details see Section 2.5.2. The general settings mentioned above closely resemble the protocols that have been employed previously by us and other groups in the context of enhanced sampling AIMD of transition metal complexes.[116,328,329] To further validate the choice of the exchange-correlation functional, the structures of the catalysts optimized in the gas phase were compared with the corresponding single x-ray crystal structure (see Tables D.1).[173] No significant difference among a set of the generalized gradient approximation (GGA) exchange-correlation functionals was found. For all simulations a cubic simulation cell with a side length of 14.56 Å, that contains the metal-oxo species of the catalyst ($[Ru^{V}O(Py5OMe)Cl]^{2+}$) as well as 107 H_2O molecules was employed. Initially, the size of the simulation cell was determined through NPT simulations at 1 bar and 300 K, subsequent equilibration at 300 K in the NVT ensemble resulted in the initial structures for the enhanced sampling calculations.

7.1.2 Bluemoon

Simulations in the *Bluemoon* ensemble were initialized from unconstrained calculations of the R or P state. The reaction path was then defined by a step-wise increase or decrease of ξ'. It turned out to be crucial that the previous simulation was sufficiently equilibrated before moving to the next step i.e. initialization of all simulations at once is not advisable. For each value of ξ' the simulations were run for about 20 ps (40'000 steps) including an equilibration time of 2.5 ps: Post-processing was performed only on the last 17.5 ps of the trajectory. Note, increasing the equilibration time to 5 ps (10'000 steps) does alter the free energies of the extrema by only $1\,\mathrm{kJ\,mol^{-1}}$ to $2\,\mathrm{kJ\,mol^{-1}}$. This is small with respect to the calculated barrier height.

The standard deviation (σ) of the average force $f_{\xi'}$ is determined by block average methods.[318] By doing so, one obtains an estimate of an upper/lower limit for the free energy difference ΔF using $(f_{\xi'} \pm \sigma_\lambda)$.[95] It is worth mentioning that this is not the only procedure to estimate the standard deviation of the thermodynamic integration of the averages force, other approaches are known in literature.[330]

7.1.3 Metadynamics

All the MetaD simulations were performed using PLUMED (version 2.4.3),[331] together with CP2K(revision 18461).[241] In order to facilitate the sampling, we used six independent walkers all of which contributed to the same metadynamics bias potential.[332] Further, the WT-MetaD formalism together with rigorous error analysis was used to enforce the convergence of the simulation.[47] Initially, Gaussians with a height of 1 kJ mol^{-1} were added to the bias potential at a pace of 50 steps (i.e. every 25 fs). The bias factor (γ) of the WT-MetaD was set to 25 in order to allow the sampling of barriers with a height of approximately 60 kJ mol^{-1}. In addition, depending on the employed CV various restraining potentials were added in order to restrict the sampling space (see description of each individual WT-MetaD simulation in Section 7.2.2). The MetaD simulations have been carried out for {Ru(Py5OMe)} employing two different sets of CVs, later referred to as 'set A' and 'set B'. Those simulations will be discussed in detail in Section 7.2.3.

7.2 Results and Discussion

In the following sections we will describe the reaction mechanism of the O–O bond formation by a WNA in detail. Employing enhanced sampling techniques such as *Bluemoon* and MetaD with various CVs we show how delicate the choice of a proper method is. In particular, the choice of an appropriate CV is a well known problem in the field of enhanced sampling and various sophisticated protocols such as the committor analysis exist in order to judge the quality of the CV.[220,333,334] Further, there have been attempts to use machine learning and artificial neural networks in order to identify appropriate CVs.[217,335] Here we do not engage in said topic, but rather follow a more chemically guided approach.

7.2.1 Bluemoon

In our previous study we have investigated a base-assisted mechanism for the O–O bond formation.[173,174] However, there can be an alternative pathway where the nucleophile releases a proton to the solvent instead of the base, the later being referred to as base-independent mechanism. Thus, several solvent molecules are required to stabilize the hydroniun or Zundel ion in proximity to the catalyst. For that reason modeling such mechanism with an approach based on geometry optimizations would have been challenging. Furthermore, a static approach would have led to serious problems regarding the exact position of the solvent molecules, proper description of the hydrogen bonding network and the localization of the corresponding extrema due to the vast increase in degrees of freedom. These limitations are alleviated in the current model system as the configurational space of both the catalyst and the solvent were explored with AIMD. Notably, under certain experimental conditions the base could have been already protonated prior to the O–O bond formation, which would in principle rule out the base-assisted mechanism.[173] However, this would result in a larger overall charge of the catalysts leading to higher reduction potentials which possibly would prevent the formation of the reactive Ru^{V}=O species. For these reasons, as well as for the sake of consistency with our previous studies, the O–O bond formation is modeled under neutral conditions i.e. none of the solvent molecules are protonated nor is the pyridine.

In line with chemical intuition the most obvious choice for modeling of the O–O bond formation by the base-assisted or base-independent mechanisms is a CV based on the distance of the metal-oxo ligand and a specific water molecule d(O_{oxo}–O_w). The two protons of the nucleophile (w) are referred to as H_a and H_b. In the case both protons are addressed at the same time the abbreviation H_ab is used. Free energy profiles obtained by the *Bluemoon* methodology employing a d(O_{oxo}–O_w) CV are shown in Figure 7.2. Time series of the force acting on the constraint, as well as a linearly interpolated force profile are shown in Figures D.2, D.3 and D.4 in the Appendix.

The larger errors shown in the base-independent pathway (see Figure 7.2, pale purple line) are partly due to non-uniform sampling of the configurational space of the nucleophile with respect to the base. To further understand this effect we would like to mention observe that most simulations describing intermediates between the TS and the R state explore configurations where the minimal distance between the base and the protons of the nucleophile, (min(d(N–H_{ab}))), is larger than 3 Å indicating weak to non existent hydrogen bonding between the nucleophile and the base (see Figure 7.3). However, the simulation where d(O_{oxo}–O_w) is constrained to 2.34 Å, explores primarily min(d(N–H_{ab})) distances around 1.70 Åsuggesting hydrogen bonding with the base.

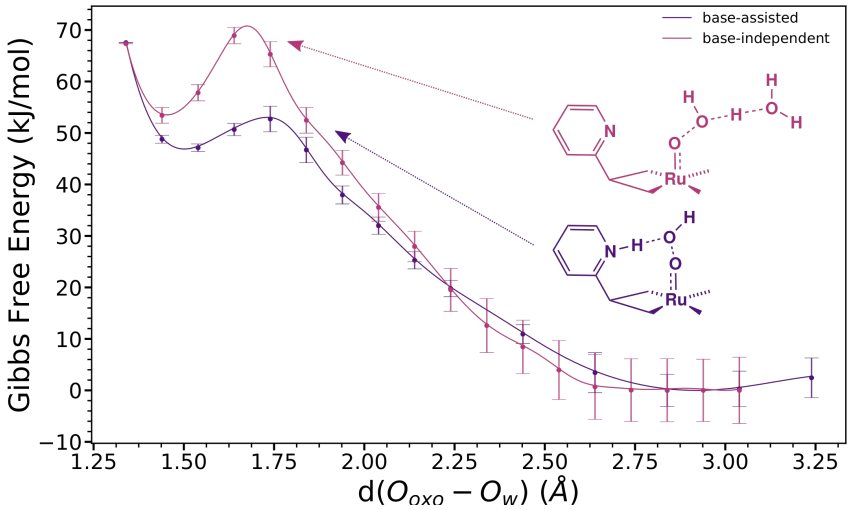


Figure 7.2: Comparison of the free energy profile for the O–O bond formation by either the base-assisted or the base-independent mechanism. Note that the increase of the error along the reaction path is caused by the integration procedure, where, by convention, the first integration point is set to 0.

Direct comparison of the two reaction mechanisms in terms of the extrema of interest i.e. R, TS and P revealed a substantially larger activation barrier of $69 \pm 8\,\mathrm{kJ\,mol^{-1}}$ in case of the base-independent mechanism as compared to the $53 \pm 4\,\mathrm{kJ\,mol^{-1}}$ of the base-assisted mechanism (see Table 7.1). This is likely a consequence of the reduced nucleophilicity of the H_2O molecule, which is destined to undergo the O–O bond formation, and is reflected by a slightly shorter O_{oxo}–O_w

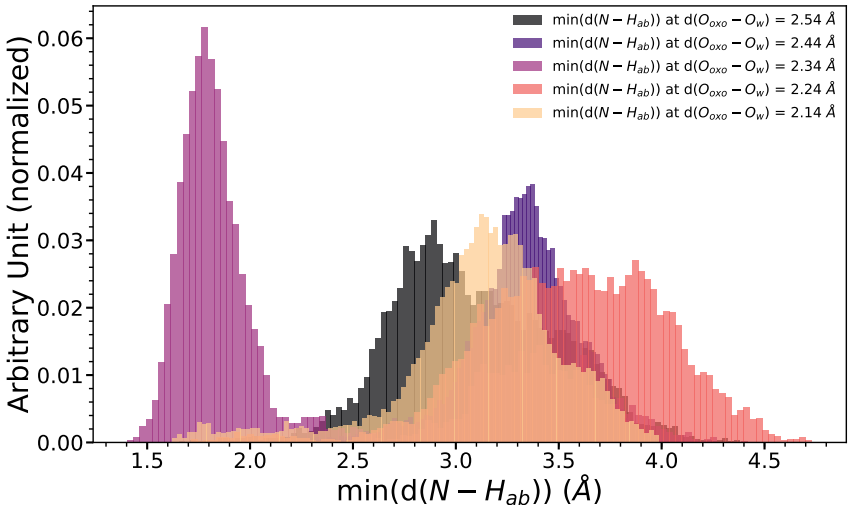


Figure 7.3: Distribution of the minimum N–H_{ab} distances evaluated for different constrained values of O_{oxo}–O_w distances indicated by the different colors for the base-independent reaction mechanism. The area enclosed by the individual histograms is normalize to unity.

distance in the TS (see Figure 7.2). Identifying the base-assisted mechanism as the energetically most favorable reaction mechanism is in accordance with our previous mechanistic proposal.[173,174]

Since multiple bonds are broken and formed in this reaction utilizing only the d(O_{oxo}–O_w) as a CV might not be a good choice, besides it being the main indication of our desired product. Visualization of the maximal O_w–H_{ab} distance as a function of d(O_{oxo}–O_w) (see Figure 7.4) revealed that in particular in the region of the transition state (d(O_{oxo}–O_w) = 1.74 Å), there is an abrupt change in the O_w–H_{ab} distance distance indicating a PT. The lack of intermediate structures describing the PT could potentially be circumvented by increasing the number of constraint simulations between 1.74 and 1.84 Å (see Figure D.5 in the Appendix). However, since the applied CV does not describe the PT explicitly, changing to a different CV is more reasonable.

To this extent, the base-assisted mechanism was modeled using a distance difference CV (d(O_{oxo}–O_w)-d(O_w–H_a)). The CV space was explored between -0.4 Å and 2.1 Å corresponding to P and R state respectively, in steps of 0.05 Å in the region of the PT and 0.1 Å elsewhere. A similar CV used by Sinha *et al.* when simulating the dehydrogenation of alcohols.[180] The obtained free energy profile qualitatively agrees well with the one obtained constraining only the O_{oxo}–O_w distance (see Figure 7.5 and Table 7.1). Nonetheless, a stabilization of both the P and the TS by 3 kJ mol^{-1} and 7 kJ mol^{-1} respectively was found, highlighting the importance of including the PT in the CV.

Inspecting the structural features of the TS revealed that in case of the distance difference CV both the O_{oxo}–O_w and the O_w–H_a distances are slightly elongated (1.78 ± 0.08 Å : 1.58 ± 0.08 Å) as compared to the distance only CV

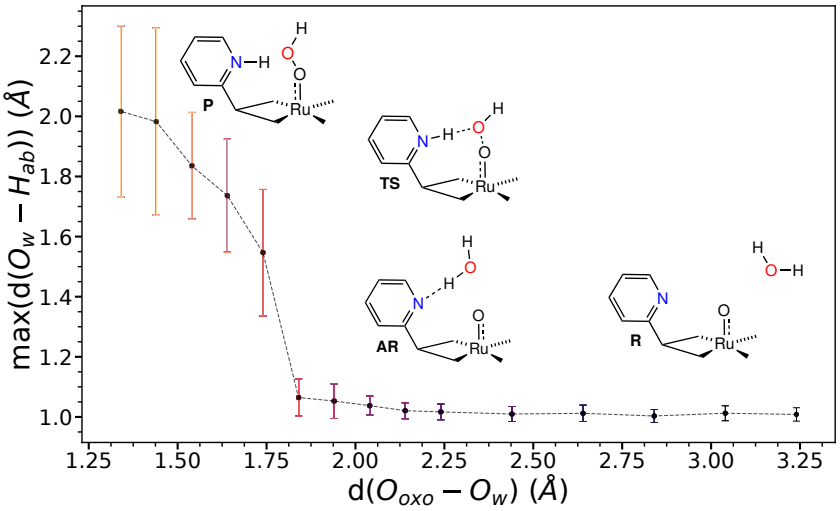


Figure 7.4: Visualization of max(d(O_w – H_{ab})) and d(O_{oxo} – O_w) distances explored for the base-assisted mechanism. The transition state corresponds to the light red point (5th from the left).

(1.74 Å : 1.5 ± 0.2 Å). This implies that the PT happens prior to the actual O–O bond formation. This is backed up by the better sampled PT, i.e. there is no abrupt PT as in the case of the distance CV only (see Figures 7.6 and D.6).

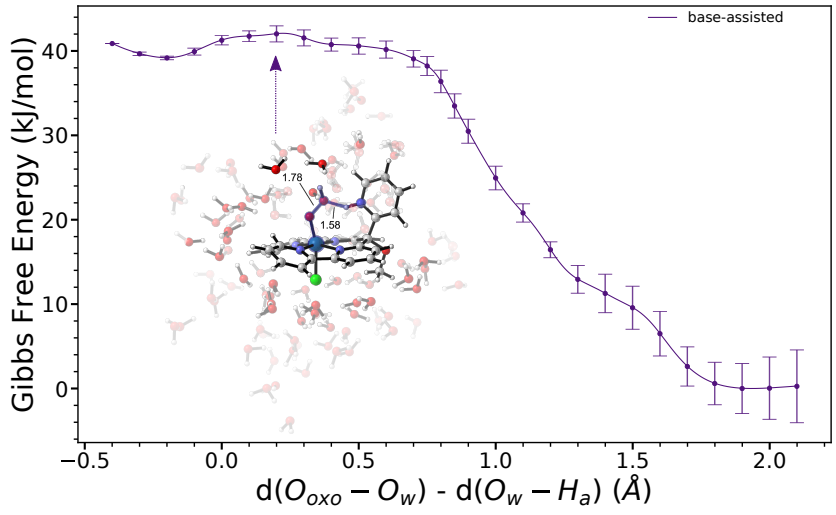


Figure 7.5: Free energy profile of the base-assisted O–O bond formation catalyzed by {Ru(Py5OMe)}, obtained by integrating the average force acting on the constraint (Equation 2.34) according to Equation 6.1. The error bars are obtained from block averaging the force profile. Inset: Representative TS structure including the solvation shell. Bond lengths are given in Å.

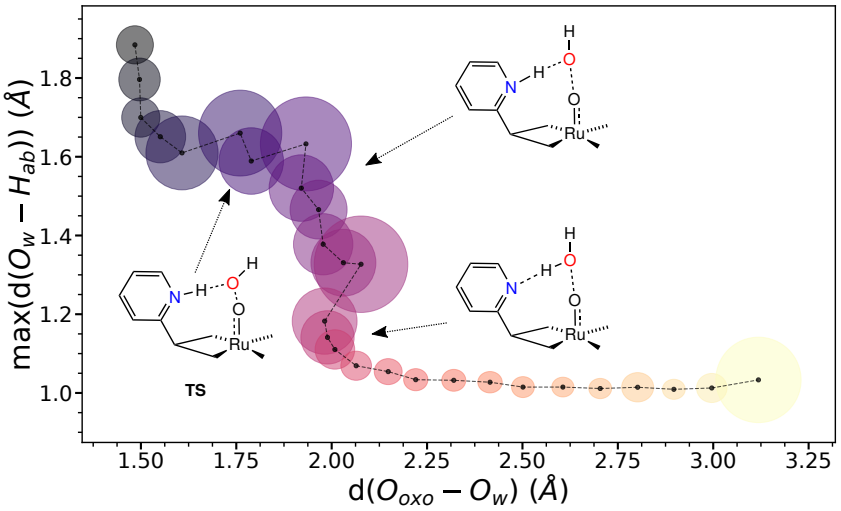


Figure 7.6: Visualization of the $O_{oxo}-O_w$ and max(O_w-H_{ab}) distances explored during the base-assisted mechanism modeled by employing the distance difference constraint. The width and height of the ellipsis represent the standard deviation of the corresponding CVs. The magnitude of it is approximately inversely proportional to the configurational space explored for each value of constrained CV i.e. a large standard deviation means that multiple combination $O_{oxo}-O_w$ and O_w–Ha distances were explored, as a consequence the number of configurations sampled for each combination alone of d($O_{oxo}-O_w$) and d(O_w-H_{ab}) is small.

The d($O_{oxo}-O_w$)-d(O–H_{ab}) CV is a better descriptor of the base-assisted O–O bond formation than the simpler $O_{oxo}-O_w$ distance only. However, it also comes with a major drawback not discussed up to now. At larger distances from the base the two protons (H_a and H_b) of the nucleophile are chemically indistinguishable. Yet, this is not reflected by the distance difference CV that enforces the release of a specific proton without controlling the spatial orientation of said proton with respect to the base. To do so in a single, chemically intuitive CV appears to be challenging. Therefore the use of a more advanced methodology, which allows the use of multiple CVs is advisable.

Table 7.1: Free energy differences between the R, TS and P highlighting the importance of the base in not only lowering the activation barrier but also in stabilizing the product with respect to the reactant. The magnitude of the standard deviation of the free energy of each individual state depends on its relative location in the reaction path, since by convention the first integration point is set to zero. The errors on the free energy differences (ΔF) are obtained by error propagation and therefore contain contributions from both ends of the reaction path. This makes the direction of the integration from R to P or from P to R less important.

	ΔF_{P-R} (kJ/mol)	ΔF_{TS-R} (kJ/mol)	ΔF_{P-TS} (kJ/mol)
$d(O_{oxo}-O_w)$ (base-assisted)	47 ± 3	53 ± 4	-6 ± 3
$d(O_{oxo}-O_w)$ (base-independent)	52 ± 8	69 ± 8	-17 ± 2
$d(O_{oxo}-O_w)$-$d(O_w-H_a)$	39 ± 3	42 ± 3	-3 ± 1

7.2.2 Metadynamics

MetaD simulations offer the option to use multiple different CVs to describe the chemical transformation of interest. Furthermore, additional degrees of freedom might be restrained to prevent the exploration of areas of the phase space that are of less importance. Studies employing either the *Bluemoon* or the MetaD methodology in the context of homogeneous water oxidation, in particular for the O–O bond formation, are scarce. In 2011, Brüssel *et al.* simulated the addition of a CO_2 molecule to a metal coordinated secondary amine, as well as the intramolecular O–O bond formation in the course of a rearrangement of a ruthenium oxo complex, employing both *Bluemoon* and MetaD simulations, where they found a simple distance CV to be insufficient to describe the intermolecular CO_2 addition, a conceptually similar observation as we made in the previous section of this work.[95] Preceding this work, Vallés-Pardo *et al.* simulated the O–O bond formation by a WNA mechanism, catalyzed by a Ru-based WOC, using the d(O–O) distance as the CV. Since their system did not possess an intramolecular base a solvent molecule acted as the base, accepting a proton from the nucleophile (here referred to as base-independent).[69] However, they did not attempt to observe multiple O–O bond formation events, instead they extracted starting points for simulations in the *Bluemoon* ensemble from the MetaD trajectory, a strategy as has been shown in the previous section, that is hardly applicable to the base-assisted WNA reaction. Piccinin *et al.* used a similar approach where the free energy surface of the O–O bond formation was explored by MetaD. However, the simulations were stopped shortly after the first occurrence of bond formation event. The energetics of the reaction were then determined by single point energy calculations employing a hybrid functional.[70] The subject of their study was a Ru-based polyoxometalate (POM), the structure of which would have allowed for an intramolecular O–O bond formation between the Ru=O and an oxygen atom of the POM cage. An initial MetaD simulation rendered the WNA energetically favorable over an intramolecular reaction. In a second MetaD simulation the WNA was modeled by employing two CVs, one to monitor the coordination number of the oxo-species and thereby the O–O distance and one to keep track of the protonation state of the nucleophile. By doing so they were able to classify the WNA as a concerted reaction consisting of the O–O bond formation and the deprotonation of the nucleophile.[70]

In the following sections we will discuss the application of MetaD simulations to

model the O–O bond formation catalyzed by {Ru(Py5OMe)}. Thereby, we will take advantage of the insights obtained from the constraint AIMD simulations. Prior to the chemical interpretation of the simulated reaction path, the convergence as well as the choice of an appropriate CV will be assessed.

Choice of the Collective Variables

The most common choices of CVs describing the O–O bond formation have been already discussed in Section 7.2.1. As has been shown, the base-assisted WNA mechanism could be described appropriately if both the O–O bond formation and the PT are taken into account. This becomes more complex since the indistinguishability of the protons has to be taken into account. In this section, a set of CVs that are able to deal with those difficulties will be introduced.

The indistinguishability of the protons can be resolved by the use of coordination numbers as CVs. Those are defined as

$$CN(r_{ij}) \;=\; \sum_{i}^{N} \frac{\left(1-(\frac{r_{ij}}{r_0})^n\right)}{\left(1-(\frac{r_{ij}}{r_0})^m\right)}, \tag{7.1}$$

where n and m define the smoothness of the switching function while r_0 specifies the turning point. The summation is carried out for a relevant subset (N) of all the atoms in the system. Usually, n and m are set to 8 respectively 16, while r_0 is system dependent. Here $r_0 = 1.5$ Å for set A and $r_0 = 1.3$ Å for set B respectively, as determined by previous trial calculations (see Tables 7.2 and 7.3).

Rather than allowing all the water molecules to act as the nucleophile, we focus the sampling on the main event of the reaction by biasing a selected water molecule. In order to assure that the selected water molecule always remained in close proximity of the reactive site, a quadratic restraining potential was introduced that prevented the $O_{oxo}-O_w$ distance to surpass 2.8 Å (see Appendix Section D.2 for more details). This distance corresponds to bulk water and was previously determined by simulations (see Section 7.2.1). By selecting a specific water as nucleophile, the definition of a second CV was facilitated as the number of indistinguishable protons was reduced to two.

We defined as the second CV (set A) the CN of the pyridyl nitrogen with respect to the nucleophile protons. Thus, we use the NH_{ab} coordination pair rather than the O_wH_{ab} CN. It is worth mentioning that the bias potential applied during the MetaD simulation depends on both CVs. Further a quadratic restraining potential (see Tables 7.2 and 7.3 for details) was used to prevent the formation of a OH^- species in the bulk of the solvent prior to the O–O bond formation, which is unlikely to exist under reaction conditions.[173]

For a second simulation (set B), an even more general description of the PT was used. The second CV in set B takes into account both the protonation state of the nucleophile and the base. This was achieved by defining it as the difference between the coordination number of the base $CN(r_{ij})$ and the nucleophile $CN(r_{kj})$:

$$\Delta CN \;=\; CN(r_{ij}) - CN(r_{kj}) \tag{7.2}$$

where r_{ij} is the distance of the solvent proton j to the intramolecular base, and r_{kj} the distance of a solvent proton j to the nucleophile. By allowing all solvent protons to be involved in the reaction, the reversibility of solvent assisted mechanisms was assured. Monitoring of the protonation state of the nucleophile has already been shown to be reasonable CV by Piccinin *et al.*[70]

Table 7.2: Overview of the MetaD simulations settings, CVs and restraining potentials for the sets of simulations A and B.

	Set A	
	description	limits
CV1	$d(O_{oxo}-O_w)$	< 2.8 Å
CV2	$CN(NH_{ab})$	-
add. restraining pot.	$d(O_w-H_a) - d(N-H_b)$ $- d(O_w-H_b) + d(N-H_a)$	-1.6 Å$< x < 1.6$ Å

Table 7.3: Overview of the MetaD simulations settings, CVs and restraining potentials for the sets of simulations A and B.

	Set B	
	description	limits
CV1	$d(O_{oxo}-O_w)$	< 2.8 Å
CV2	$CN(NH) - CN(O_wH)$	> -2.0
add. restraining pot.	-	-

Assessing Convergence

Assessing the convergence of MetaD simulations requires an in-depth analysis of the states sampled and their relative statistics.[46,220] This is in particular true in the context of AIMD where the computational cost is a severe limiting factor.

In general convergence is achieved when the simulation freely transits between the states of interest. This is equivalent to the reoccurring observation of the chemical transformation of interest. The time traces of the CVs in Figure 7.7 show multiple O−O bond formations and breakages as well as the protonation and deprotonation events of the base within the same walker and across multiple walkers. When the free energy difference between the reactant and product is not of interest, it is common to stop MetaD simulations after the first occurrence of the reaction of interest.[336] In this work, we intentionally went beyond this simplification in order to get a full picture of both the reactant and product states and their energy difference.

Another indication of convergence can be obtained by projecting the free energy surfaces (FESs) onto the degrees of freedom that were biased i.e. in the case of set B the $O_{oxo}-O_w$ distance and the difference of the coordination numbers (CN(NH)-CN(O_wH)). An estimate of the standard deviation of the free energy profile is shown in Figure 7.8. It was obtained by calculating block averages (5 blocks) over the concatenated trajectory of the six individual walkers (see also Figure D.7 in the Appendix). In general the standard deviation was found to be within a few kJ/mol indicating uniform sampling of the whole CV space. Note the variable magnitude of

error over the CV space indicates different degrees of convergence. This is especially true for the upper and lower limit of the explored CV space where the sampling is poor.

The convergence was further validated by calculating an estimated error on the MEPs connecting the R and P state (see Section 7.2.4). Applying block averaging methods an optimal block number of 5 was determined for the error on the MEP (see Figure D.8 in the Appendix). Note the standard deviations given here only account for the statistical error of the sampling, not the systematical errors of DFT caused by e.g. the choice of the exchange-correlation functional and basis-set.

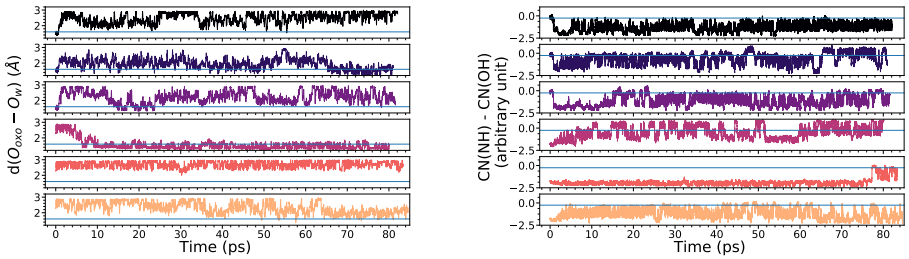


Figure 7.7: Left: Time-trace of the O–O distance for different walkers of {Ru(Py5OMe)} (set B). The blue horizontal line indicates the formation of the O–O bond. Right: Time-trace of CN(NH) - CN(OH) of the same simulation. Formation of the hydroperoxo species is indicated by the horizontal blue line. Note the horizontal lines mere serve as a visual guide line rather than a strict assignment of states.

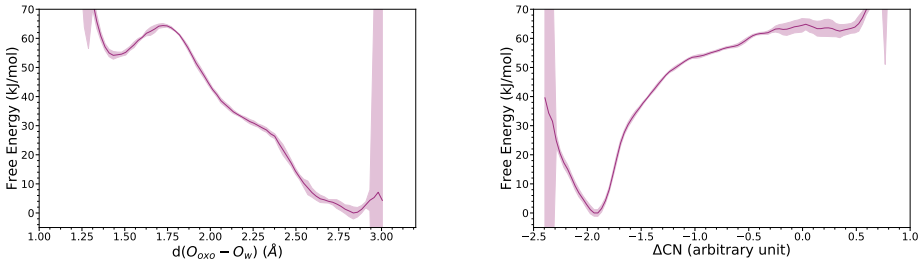


Figure 7.8: Left: Free energy profile including the error range of the $O_{oxo}-O_w$distance for {Ru(Py5OMe)} (set B). Right: Free energy profile including the error range of the CN(NH) − CN(O_wH) CV of the same simulation. Note the free energy profiles shown here were obtained by reweighting procedures accounting for the additional bias imposed by the restraining potentials.

7.2.3 Collective Variable Analysis

The CVs biased in set A and set B both successfully describe the O–O bond formation by a base-assisted mechanism, which can be seen by the exploration of the transition state region (set A: CV1 ≈ 1.8 Å, CV2 ≈ 0.9 and set B: CV1 ≈ 1.7 Å, CV2 ≈ -0.2) (see Figure 7.9). There were significant differences with respect to both reaction pathways and states explored between the two different sets. This can be

seen in Figure 7.9 where the simulations biasing the CVs from set A were reweighted according to the CVs from set B and vice versa.

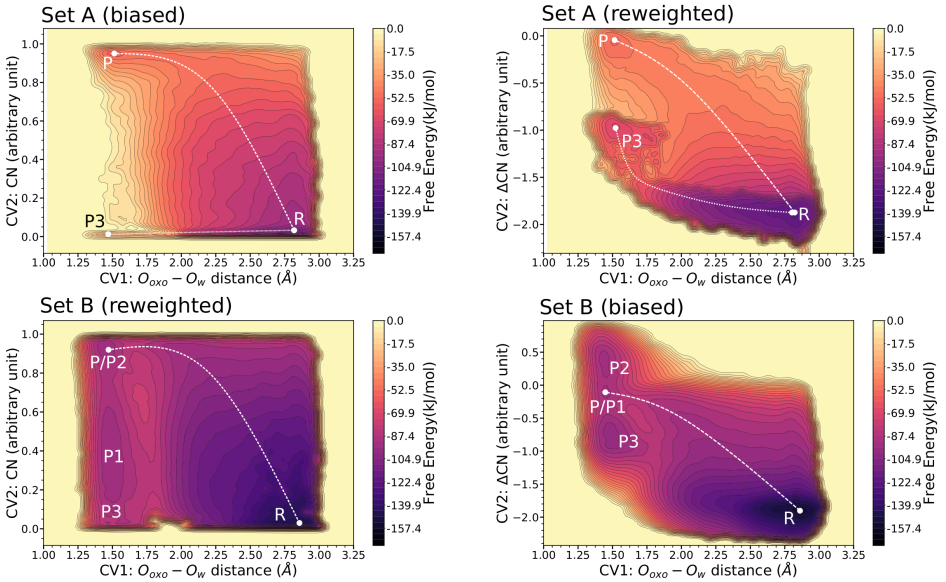


Figure 7.9: Top-left: FES of {Ru(Py5OMe)} reconstructed using the settings of set A. Top-right: Reweighted FES (set A) according to CVs of set B. Bottom-left: Reweighted FES (set B) of {Ru(Py5OMe)} according to CVs of set A. Bottom-right: FES of {Ru(Py5OMe)} reconstructed using the settings of set B. The dashed line is a visual aid to indicate the base-assisted pathway, analogously the base-independent is indicated by the dotted line.

The first notable difference amongst the sets of CVs is that the base-independent reaction path, directly connecting the R and P3 state (see Figure 7.10), only occurs when the CVs from set A are biased by MetaD, which is a key difference between the two simulations (see Figure 7.9, top-left and top-right panel). Thereby, a proton is released to the nearest solvent molecule and from there transfered by the Grotthuss mechanism to different solvent molecules. In the case of set A, the protons to be transfered to the base are explicitly defined. Due to this, if one of those protons hops to the solution there is no direct bias bringing it back to the reaction region, consequently the reaction halts and the simulation is then stuck into sampling product like states, which compromises the overall statistics. The possibility of a base-independent reaction in the case of set B will be discussed in detail in Section 7.2.4.

The second striking difference is the co-existence of three energetically similar local minima in which the O−O bond has been formed in case of the simulation employing the CVs of set B (see Figure 7.9, bottom-right panel). Those are states with either a formed hydroperoxo species and the protonated base (P), a peroxo species and a protonated base where a proton was released to the solvent (P2) or a hydroperoxo species with a deprotonated base where the proton of the nucleophile was released to the solvent (P3) (see Figure 7.10). The P3 state corresponds to the product of the base-independent pathway, but can also be formed via PT from the hydroperoxo ligand of P to the solvent or by the deprotonation of its base. A possible intermediate

structure where a proton is shared between the base and the peroxo ligand (P1) is indistinguishable from the P state within the CVs from set B. The absence of those species in the case of set A is likely a consequence of the applied restraining potential on the $d(O_w - H_{ab})$ and $d(N - H_{ab})$ distances. This potential was introduced based on previous trial runs in order to keep the 'active' protons (H_a and H_b) in between the base and the nucleophile.

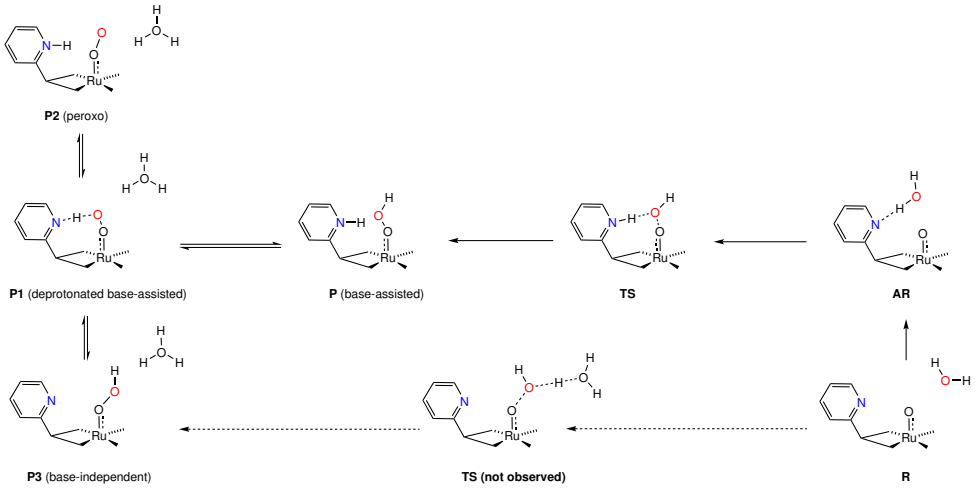


Figure 7.10: Schematic overview of the reaction network connecting the observed states. Note that P and P1 co-inhabit the same free energy basin and thus are virtually indistinguishable within the chosen CVs.

7.2.4 Characteristics of the FES

MEPs were identified on the smoothed FES using the minimum energy pathway analysis for energy landscapes (MEPSA) employing a variation of Dijkstra's algorithm.[186,337] MEPs obtained by this procedure by no means have an exclusive character i.e. they are the lowest energy pathways connecting two predefined local minima, but there might be several alternative pathways that are energetically similar. By evaluating the final MEP on the FESs obtained for the individual blocks used for the block analysis an estimate of its error can be given (see Figure 7.13).

The inherent complexity of this process translate into an equally intricate FES making the unambiguous identification of the MEP difficult (see Figure 7.11). In order to differentiate the two pathways, the FES was projected into a third dimension (see Figure 7.12). As shown in Figure 7.9, following the formation of the N–H bond can directly monitor the contribution of the base to the overall reaction, thus identifying a clear separation amongst the two possible reaction pathways. Conceptually similar representations have been used in cases where more than two CVs were required to model the reaction.[338] The FESs clearly shows how the local minima representing the R and the P state are connected trough a low energy transition region, highlighting the fact that the transition region was exhaustively explored during the simulation. On the other hand, there is no low energy region connecting the R and the P3 state as it would be required in the case of the base-independent pathway. The three dimensional representation of the FES in Figure 7.12 further implies that, as soon as

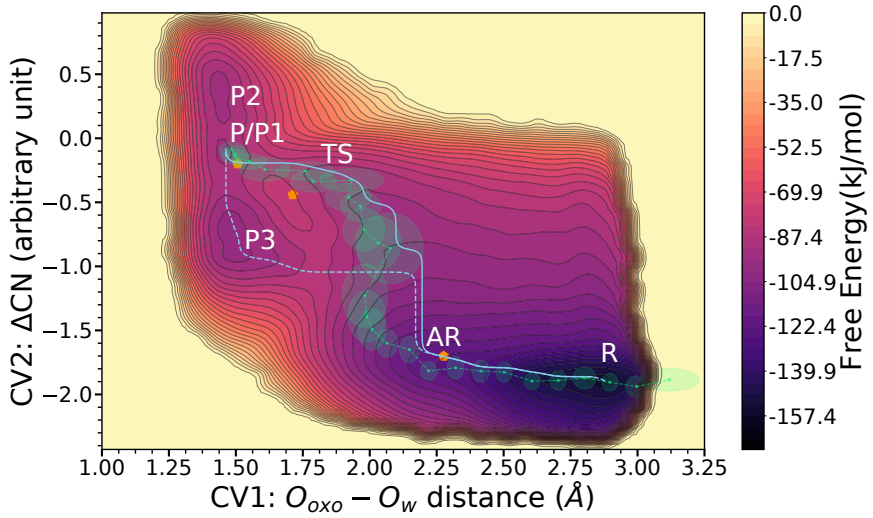


Figure 7.11: Visualization of the MEP for the O−O formation on the FES of {Ru(Py5OMe)} (set B) following the base-assisted mechanism (solid line) and the base-independent mechanism (dashed line). The base-assisted reaction path sampled within the *Bluemoon* ensemble projected onto the same FES is represented by the pale green dashed line and the ellipses. The height and width of the ellipses corresponds to the standard deviation of the CVs calculated from the trajectories of the constrained simulations. The orange pentagons correspond to the AR, TS and P state obtained by the static DFT simulations reported by Schilling *et al.*[174]

P is formed, PT from the base to the hydroperoxo ligand and from the hydroperoxo to the solvent ligand becomes energetically feasible. Therefore, the formation of the O−O bond without the involvement of the base, as suggested by the dashed-MEP (see Figure 7.11) could be excluded leaving the base-assisted MEP as the only viable option. This is in agreement with the *Bluemoon* simulations discussed in Section 7.2.1. A set of exemplary structures along the preferred base-assisted MEP is shown in Figure 7.13. During the first phase of the reaction, the nucleophile is in close proximity to the catalyst, then it gets in a proper orientation for the O−O bond formation by engaging in hydrogen-bonding with the intramolecular base. Upon a successful PT, the distance between the nucleophile and oxo-ligand is reduced, leading the proposed TS. Continuing from there the O−O bond is then finally formed. During the whole reaction several solvent molecules engage in hydrogen-bonding with the nucleophile. This is particularly pronounced for intermediate states where the proton is shared between the nucleophile and the base.

The validity of the base-assisted MEP MetaD biasing the CVs from set B was further verified by comparing it with the MEP obtained on the reweighted FES (biased CVs from set B, but reweighted according to the set A CVs). The free energy differences between the extrema (R, TS and P) of those simulations are given in Table 7.4. For the MetaD simulation biasing set B, as well for set B reweighted according to set A the relative energies are qualitatively the same. Naturally, reweighting results in larger errors, in particular for the reactant state which due to the generalized definition in set B (i.e. the coordination number of all protons instead of two specific

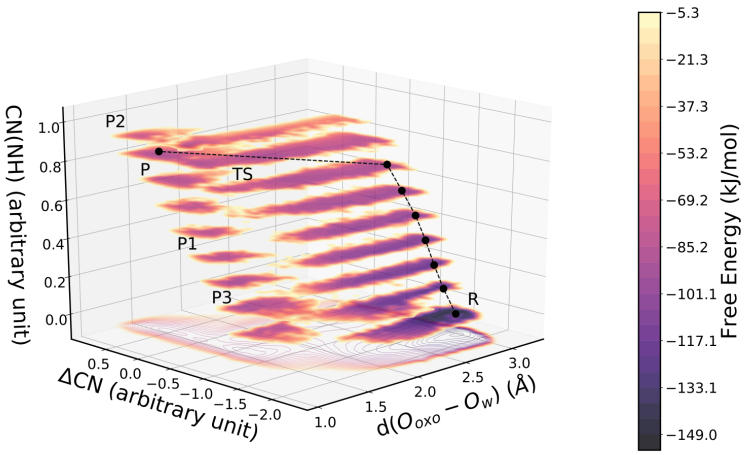


Figure 7.12: FES of {Ru(Py5OMe)} reconstructed using the MetaD simulations with the set B of CVs (x- and y-axis). Additionally, a third dimension, the coordination number of the base (CN(NH)) (z-axis), is used to visualized the the preferred reaction path. Note that the path shown in the graph only serves as a visual aid and does only qualitatively correspond to the MEP shown previously.

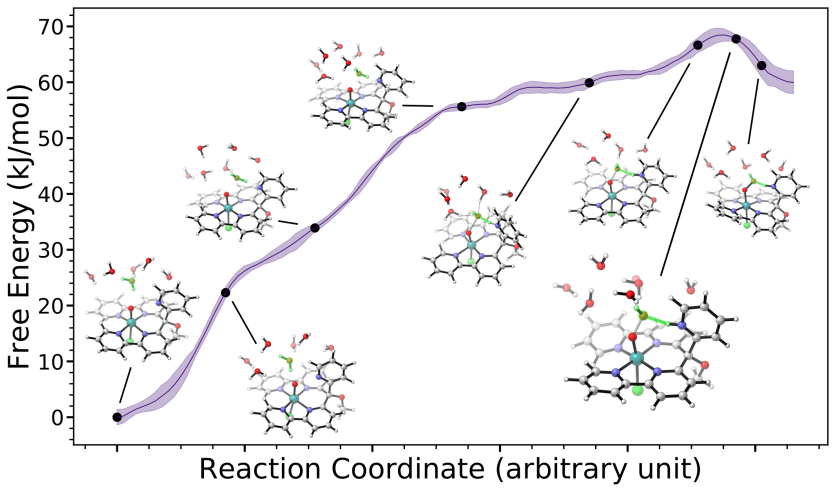


Figure 7.13: MEP of the base-assisted O–O bond formation of {Ru(Py5OMe)} obtained from the MetaD simulations with the set B of CVs. An estimation of the error of the MEP is shown together with exemplary structures along the reaction coordinate. Those were obtained by the clustering algorithm implemented in METAGUI 3.0.[339]

ones) is slightly different among the two sets of CVs. The larger errors associated with set A highlight the fact that the simulation did not reach the same sampling

efficiency as the one of set B, especially due to the reversibility of the hydronium formation as discussed in Section 7.2.3. Note, since the simulation employing the CVs of set A got stuck, its length is only about half of the simulation using the CVs of set B. Nonetheless, qualitatively stability of the reactant, transition state and product are recovered.

Table 7.4: Free energy differences obtained from the MEP of the FES. The reweighted pairs, highlight the difference sampling efficiency amongst the CVs sets. All values given are in kJ/mol.

	{Ru(Py5OMe)}		
	ΔF_{P-R}	ΔF_{TS-R}	ΔF_{P-TS}
MEP (set A)	59 ± 48	64 ± 42	-6 ± 24
MEP (set A, reweighted B)	70 ± 21	77 ± 17	-7 ± 26
MEP (set B)	60 ± 3	69 ± 2	-8 ± 3
MEP (set B, reweighted A)	55 ± 9	65 ± 9	-10 ± 3

7.2.5 Confronting Sampling Methodologies

In our previous study, we applied static DFT simulations to model the O−O bond formation, whereby our model system contained two explicit water molecules, one of which acted as the nucleophile. In the TS structure the additional water molecules engaged in hydrogen bonding with the nucleophile locking it in place for the O−O bond formation. Lack of additional solvent molecules with whom hydrogen bonding would have been possible resulted in the AR structure that closely resembled the TS. As a consequence, the energetics of the AR, TS and P were referenced against the corresponding Ru^{V}=O species plus two H_2O, obtained by separate optimizations, as they better represent the energetics of the R state.[173,174] Projecting the AR, TS and P states reported in the earlier study onto the FES obtained by the MetaD simulations (see Figure 7.11) reveals that neither the AR nor the P are local minima.[174] This comes as no surprise since the underlying electronic structure method as well as the description of the surrounding solvent are different from the current MetaD simulations. A detailed comparison of key structural features of the extrema describing the O−O bond formation can be found in Section D.4 in the Appendix. In the case of the static DFT simulations geometries were optimized at the BP86-D3/def2-TZVP level of theory within the TURBOMOLE program package.[206,224–226] The reported free energies were then obtained using single point calculations at the B3LYP-D3/def2-TZVP and the rigid rotator harmonic oscillator approach, whereby the interaction with the solvent were approximated by COSMO. Those values were further corrected for ambient temperature and pressure (see Schilling *et al.* for more details[174]).[197,198] Nevertheless, the structures are similar, and at that level of theory one might thus gather valuable qualitative insights in case of obvious spurious agreements among the methods. In this specific case, the structure of the TS and P are located in close proximity of their correspondent counterparts on the FES obtained by MetaD simulations. Further, the AR lies approximately on the MEP connecting the R and P on the FES and is clearly different from bulk water supporting our choice of referencing the free energies against the Ru^{V}=O species.

When evaluating the free energies of the AR, TS, and P states obtained by static

calculations on the FES a surprising agreement between the free energies of the TS was found, despite the vast differences in the applied methodologies (see Table 7.5). On the other hand, substantial differences in the stability of the P state were found, that are likely attributed to the fact that the static DFT structures are locked in a configuration by hydrogen bonding between the OOH ligand and the one additional solvent molecule (d(O_{HOO}–H) = 1.69 Å (static) vs. 2.0 ± 0.2 Å (MetaD)). Further, the static solvation did not allow the exploration of different protonation states such as the deprotonated product or the peroxide species.

When describing the reactant using DFT geometry optimization, weak interactions between the base and the metal-oxo might additionally stabilize the reactant state. This interaction is best described by the d(N–O_{oxo} = 2.4 Å) distance and the angle of the pyridyl with respect to the plan spanned by of the Ru metal center and the ligand ∠($C_{py}C_{OMe}$Ru) = 100.1 ° (see Figure D.1 in the Appendix). Within the enhanced sampling methods used in this work, this distance increases to d(N–O_{oxo} = 2.7 ± 0.3 Å and ∠($C_{py}C_{OMe}$Ru) to 104 ± 5 ° implying an even weaker interaction between the Ru and the pyridyl but no complete absence of interactions. This further suggests that there is no spontaneous oxygen atom transfer from the metal to the pyridyl forming a pyridyl-N-oxide, rendering such a reaction pathway less likely.

Table 7.5: Free energies of the AR, TS and P structures reported by Gil-Sepulcre *et al.* and their corresponding relative value when projected onto the FES (set B). The free energies of the MetaD simulations are given relative to the local minima corresponding to bulk water. All free energies are given in kJ/mol.[173]

	{Ru(Py5OMe)}			
	R	AR	TS	P
static DFT[174]	0	29	65	26
static DFT projected on FES	0 ± 2	31 ± 2	71 ± 2	63 ± 3
MetaD (set B)	0 ± 2	-	69 ± 2	62 ± 3

Analogously, the reaction pathways obtained by the *Bluemoon* simulations using the difference of distances constraint and the MetaD simulations were compared by projecting the respective MEPs onto the FES (set B) (see Figure 7.11). The reaction path explored by the *Bluemoon* simulation follows the same features of the FES as the MEP, thereby visiting similar extrema i.e. the R, the TS and the P. Consequentially, the free energies of the projected extrema were almost identical to the extrema visited by the MEP (see Table 7.6). A direct comparison of the relative stabilities of extrema suggested that within the *Bluemoon* ensemble the reactant state is destabilized by almost 20 kJ compared to the MetaD simulation. The difference is likely a consequence of the rather simple CVs used within the *Bluemoon* ensemble. The mean force therefore lacks significant contributions from other degrees of freedom relevant for the O–O bond formation.[95] Those other degrees of freedom appear to be in particular important for the description of the PT. A further limitation of the CV arising from the explicit specification of the 'active' proton was already discussed in Section 7.2.1.

Table 7.6: Comparison the energetics of the R, TS and P states obtained with *Bluemoon*, MetaD and the projection of the *Bluemoon* path onto the MetaD FES. All free energies are in kJ/mol.

	{Ru(Py5OMe)}		
	R	TS	P
Bluemoon	0 ± 2	42 ± 3	39 ± 3
Bluemoon path projected on FES	0 ± 2	69 ± 3	61 ± 3
MetaD (set B)	0 ± 2	69 ± 2	62 ± 3

7.3 Summary and Conclusion

Applying forefront *ab initio* molecular dynamics in combination with enhanced sampling methods we investigated in detail the O−O bond formation catalyzed by the {Ru(Py5OMe)} WOC. Using a large number of explicit solvent molecules in a periodic framework treated at ambient conditions and at the DFT level allows for a sophisticated description of the solvent and finite temperature effects and is thus a comprehensible approach on how to improve the picture obtained by static DFT calculations. We showed that a base-assisted mechanism is the energetically favorable pathway for the O−O bond formation by a WNA contrary to a base-independent mechanism. This was achieved in two ways, first by forcing the exploration of different pathways within the *Bluemoon* ensemble, and secondly by sampling pathways in an unrestrained manner utilizing the MetaD simulations. In particular, modeling of the base-independent mechanism would have been challenging within a static solvation approach as used in previous work, because the stabilization of a hydronium in proximity of the catalyst would require several additional solvent molecules. The inclusion of which would require the use of sampling techniques otherwise the results would become ambiguous.

Focusing on the base-assisted mechanism we have shown how the interplay between the base and the nucleophile results in the formation of a hydrogen-bonded reactive intermediate prior to the O−O bond formation. A proton transfered to the base subsequently leads to the formation of the O−O bond. The refinement of the CVs choice used to monitor the O−O bond formation demonstrated that, being in line with chemical intuition, simple reaction coordinates can lead to the exploration of unexpected areas of the FES, but may miss relevant features of the process of interest.

A direct comparison between energetics of the reaction path simulated by either the static, *Bluemoon* or MetaD is usually not expedient due to the differences in the methods. However, there was still a remarkable agreement among the structures of the local extrema (R, TS, and P). The energetics obtained in that manner were in agreement among the three different methodologies. To the best of our knowledge this study is the first direct comparison of those different simulation protocols in the context of a O−O bond formation by a WNA and is therefore important for the emerging field of water splitting for sustainable energy storage and conversion.

The biggest qualitative difference between energetics obtained by the static DFT simulations and the ones obtained by enhanced sampling AIMDs were related to the product state, which turned out to be less stable than suggested by the optimizations. On the other hand, the instability led to a variety of product states that

were isoenergetic, among them there were deprotonated species which represent expected intermediates further down the catalytic cycle, such as the peroxo-species (P1) suggesting a high reactivity of the P. The detection of such isoenergetic states are a clear indication of how necessary a more robust sampling is, since it can unveil new system capabilities which can be paramount in other applications.

Overall, we have shown that modeling of the O–O bond formation by a WNA within a static DFT framework and asserted limitations due to the description of the solvent dynamics can lead to an oversimplified picture of the reaction of interest and the reaction network as well as potentially wrong estimates of the reaction barrier. Enhanced sampling methods such as *Bluemoon* and MetaD offer a suitable solution to proper sample solvent dynamics and finite temperature effects while having many explicit solvent molecules. However, they come at a significant cost. Nevertheless, this then allows unprecedented insight in the process studied and the relevant parts of the phase space.

Chapter 8

Summary and Outlook

In the future, artificial water splitting will play a major part in the energy supply chains. However, to fully reveal its potential, in-depth knowledge of the underlying reaction mechanisms is of great importance. In this work, we have shown our contribution to this crucial task at hand.

In order to set the stage for the chapters to follow, we first introduce the generally accepted mechanistic proposals of transition metal catalyzed water oxidation. We paid special attention to the O–O bond formation, that is often thought to be the bottleneck of catalytic water oxidation. The mechanistic proposals and their respective intermediates were then linked to various computational methods and simulation protocols with the aim to showcase some of the methodical and technical opportunities as well as challenges that arise when modeling complex systems.

In the first part of the work, we focused on a selection of Co-based WOCs. Besides giving a comprehensive overview on the mechanistic proposals, we also gave an abstract of the studies on Co^{II}-based cubanes developed by Patzke and co-workers. Among them are unique lanthanide containing cubanes that mimic nature's OEC. Simulations of those WOCs revealed an unusual structural flexibility which manifests itself in the existence of an 'open' and 'closed' cubane core structure. In an exhaustive study, we showed that those 'open' conformers are indeed thermodynamically stable and kinetically accessible, but they do not enhance the catalytic performance. On the contrary, it became clear that extensive explicit static solvation biases the thermodynamics due to a lack of conformational sampling.

Consequentially, the third generation of the Co^{II}-based cubanes was modeled employing explicit dynamic solvation by means of AIMD simulations. The consideration of ambient conditions and directed solvent-solute interactions not only revealed the high flexibility of the cubane core, but also highlighted the non-innocent character of ligands which were found to participate in proton-reshuffling among the various protic groups. This adds an additional layer of complexity to a system for which, due to its unique alignment of the aqua-ligands, a large number of possible reaction pathways are plausible.

In collaboration with our experimental co-workers in the groups of Alberto and Llobet, we focused our attention on mononuclear Ru-based WOCs. In a first study, we compared the O–O bond formation catalyzed by a bda-based WOC via either a WNA or a RC mechanism. In the case of pic derivatives of the WOC, the activation energies were found to favor the WNA mechanism. This is in agreement with the

first order kinetics observed in the experiments. On the other hand, in the case of the isoq derivative neither the experimentally determined rate order nor the simulations were fully conclusive in terms of a dominant a reaction mechanism. This may be due to various reasons such as solvation effects, the spatial orientation of the WOC as well as experimental uncertainties. There is also experimental evidence for catalyst decomposition.

Moving to related Ru-based WOCs bearing the Py5 ligand, we showed that contrary to an initial hypothesis based experiments, a small modification to a non-coordinating region of the ligand did neither alter the thermodynamics nor the kinetics i.e the heights of the activation barriers of the proposed water oxidation mechanism. This observation triggered a series of experiments which together with our simulations suggested that the difference in the catalytic activity might be caused by the quicker deactivation of the catalyst, a fact that is often overlooked. Having studied the Py5 WOCs in-depth, we proposed modifications to the dangling base in order to improve WOCs catalytic performance. Altering its basicity turned out to be a key factor in order to lower the activation barrier of WNA. In this sense, the study is a first step towards *in silico* design. The simulation of the TSs of those reactions required the inclusion of an additional solvent molecule which raised several questions regarding the nature of the reactant state as well as the stabilization of the TS. Again, the inclusion of explicit solvation appeared to be necessary. As a first step, we showed that the *Bluemoon* method can be used to reliably predict pK_a values of complex systems such as the Ru-based WOC. To do so, we modeled the pK_a values of the Ru-WOCs submerged in a box of more than a hundred water molecules. We benchmarked three different protocols to calculate pK_a values and showed that they are all able to reproduce experimental trends but diverge in predicting the values quantitatively. Our study is the first of its kind when it comes to directly comparing different simulation protocols and thereby serves as a guideline for the application of the *Bluemoon* method to determine pK_a values of transition metal complexes as WOCs.

In the latest study, we applied the *Bluemoon* method as well as WT-MetaD in order to model the O−O bond formation catalyzed by the Py5 WOC. This did not only allow us to confirm our previous mechanistic proposal for the O−O bond formation i.e. a base-assisted WNA, but we gained valuable insights on the underlying reaction network. In particular, the product state turned out to undergo proton transfer reactions between the base, the hydroperoxo ligand and the solvent. Exploration of those states would not have been possible using the implicit solvation approach used in previous studies. In our study, we showcased that only properly selected CVs are able to fully capture the phase space relevant to O−O bond formation. We also provide a comprehensive guide on how to tailor the CV to the system at hand. Having simulated the O−O bond formation in a static framework employing geometry optimizations and implicit solvation as well as in a fully dynamic manner by the *Bluemoon* and WT-MetaD methods allowed us to compare those methodologies in terms of the structures of the reactant, transition and product state. In general the structures were found to be similar among the different methods. However, in particular in the case of the product state, the dynamic treatment of the solvent led to structures that exhibited a lesser degree of hydrogen bonding between the catalyst and the solvent. This resulted in more conformational freedom of the hydroperoxo ligand than in the case of the static calculations. A similar observation can also be

made when comparing the structure associated reactant with the one of the reactant resembling the catalyst and bulk water.

Throughout this work we have introduced several WOCs, which were studied by a broad range of methods and simulation protocols. In most cases, our results were in good agreement with the experimentally determined observables given the many explicit and implicit assumptions and uncertainties in the calculations and the experiment. The bottom line of all those studies is that inclusion of explicit solvation is advisable if the reaction of interest i.e. the O−O bond formation, takes place by an intermolecular reaction such as a WNA or a RC. Initially, this might be done by static calculations in combination with implicit solvation models. Yet, much more insight will be obtained with dynamic approaches and the application of enhanced sampling methods. In this work, this has been accomplished by combining forefront DFT-MD simulations of systems containing a solute and a large number of solvent molecules with enhanced sampling techniques such as MetaD and *Bluemoon*.

No matter by which protocols the catalytic systems are modeled, reality will always be much more complex. Catalytic solutions usually contain a large variety of compounds i.e. counter ions, buffer molecules, sacrificial electron acceptor etc. all of which affect its ionic strength and pH value. Those parameters are usually easy to control in experiments but including them in calculations is difficult. This is why they are often neglected, what naturally leads to some discrepancies between experiment and simulation. Furthermore, experiments are rarely able to unambiguously identify the exact structure of the catalytically active species. This can lead to an additional layer of uncertainties when it comes to the interpretation of measured TOFs and the deduction of an underlying reaction mechanism. An appropriate structural model together with the in-depth understanding of all the environmental parameters relevant to catalytic activity are the premises on which computational chemist build their model systems. Unfortunately, the information required to do so is often hard to come by. However, computational chemistry offers the opportunity to use simplified model systems where certain parameters can be varied selectively. The identification and quantification of the effects those parameters have on the model system, can lead to design guidelines on how to fine-tune certain properties like the activation barrier of a reaction. Those guidelines are the first step of *in silico* design.

The WOCs studied in this work still pose many unresolved riddles. In particular, the Co^{II}-based cubanes are a suitable target for follow-up studies. Unraveling the mechanism of the O−O bond formation for those clusters would be highly desirable and are in principle possible by applying the same tools as for the Ru-based catalysts studied in this work. However, the combinatorial challenges are immense and with today's available computational resources hard to achieve by AIMD. On the other hand, DFT or even wavefunction based methods are required to accurately describe those clusters. A practicable solution to the problem might be to separate questions that require a combinatorial approach from the ones that require a highly accurate electronic structure method.

Also the Ru-based WOCs offer a great potential for further studies. Be it exploring the feasibility of 'exotic' O−O bond formation mechanism involving a pyridyl-N-oxide or by continuing the *in silico* design pathway by also modifying the bipyridyl subunits. Then there are also other aspects of the water oxidation cycle that deserve more attention such as the dissociation of O_2 or the association of the substrate.

In our work, we demonstrate how modeling of thermodynamics and kinetics can be used to guide experimental studies and in some cases even lead to design guidelines for better WOCs. Our studies were conducted employing methods and simulation protocols that are at the forefront of the current research in the field of transition metal catalyzed water oxidation and may serve as a template for further studies in this field.

Chapter 9

Appendix

A Discovery of open cubane-core Structures

The relative free energies ΔG_{Si-S0} were compared to the computational limit of water oxidation, i. e. the reaction free energy of water oxidation calculated with the same methods as the free energies of the catalytic state (108.2 kcal mol^{-1} for the CP2K calculations, and 115.5 kcal mol^{-1} for TURBOMOLE calculations using COSMO). The relative free energies were scaled by the ratio of the computational and experimental free energy of water oxidation, i.e. 113.5 kcal mol^{-1} / 108.2 kcal mol^{-1} = 1.05 (CP2K) and 113.5 kcal mol^{-1} / 115.5 kcal mol^{-1} = 0.98 (TURBOMOLE).

A.1 Different Solvation Models

Table A.1: Relative free energies calculated for $\{Co_3Er(pyMeO)_4\}$ with explicit solvation shell; Free energy differences are given in kcal mol^{-1}. [a]: Including the zero-point corrected free energy difference between S0 (closed-open)

State Si	S0	S1	S2	S3	S4
ΔG_{Si-S0} (closed)	0.0	37.7	72.5	96.7	120.5
ΔG_{Si-S0} (open)	0.0	36.1	64.3	114.5	138.9
ΔG_{Si-S0} (closed) (no solvent)	0.0	19.9	57.6	88.6	117.3
ΔG_{Si-S0} (open) (no solvent)	0.0	23.3	85.3	96.7	131.4
$\Delta(\Delta G_{Si-S0})^{[a]}_{closed-open}$	18.0	19.6	25.9	0.7	0.2
$\Delta(\Delta G_{Si-S0})^{[a]}_{closed-open}$ (no solvent)	8.5	5.3	-18.3	0.7	-5.2

Table A.2: Relative free energies calculated for $\{Co_3Tm(pyMeO)_4\}$ with explicit solvation shell; Free energy differences are given in kcal mol^{-1}. [a]: Including the zero-point corrected free energy difference between S0 (closed-open)

State Si	S0	S1	S2	S3	S4
ΔG_{Si-S0} (closed)	0.0	39.4	61.1	117.8	114.8
ΔG_{Si-S0} (open)	0.0	34.8	53.6	103.5	122.3
ΔG_{Si-S0} (closed) (no solvent)	0.0	28.1	30.4	84.9	110.2
ΔG_{Si-S0} (open) (no solvent)	0.0	28.4	77.9	121.7	130.8
$\Delta(\Delta G_{Si-S0})^{[a]}_{closed-open}$	6.3	10.8	15.5	20.1	28.1
$\Delta(\Delta G_{Si-S0})^{[a]}_{closed-open}$ (no solvent)	12.2	11.9	-4.7	-23.4	-7.8

Table A.3: Relative electronic energies calculated in TURBOMOLE between the catalyst structures optimized (opt) in TURBOMOLE using COSMO and without explicit solvent molecules the electronic energies of which were obtained from a single point calculation with Turbomole and COSMO (sp) on the catalyst structure of the geometry optimization including explicit solvation shell; Electronic energy differences are given in kkcal mol^{-1}.

State Si	S0	S1	S2	S3	S4
$\Delta E_{Si(opt)-Si(sp)}$ $\{Co_3Er(pyMeO)_4\}$ (closed)	-8.3	-10.4	-2.2	-15.3	-19.2
$\Delta E_{Si(opt)-Si(sp)}$ $\{Co_3Er(pyMeO)_4\}$ (open)	-7.8	-4.4	-13.9	-0.3	-1.3
$\Delta E_{Si(opt)-Si(sp)}$ $\{Co_3Tm(pyMeO)_4\}$ (closed)	-26.0	-31.5	-30.1	-22.7	-23.8
$\Delta E_{Si(opt)-Si(sp)}$ $\{Co_3Tm(pyMeO)_4\}$ (open)	-20.2	-18.9	-11.8	-16.6	-12.0

Table A.4: Relative free energies calculated for $\{Co_3Er(pyMeO)_4\}$ employing either COSMO or DCOSMO-RS; Free energy differences are given in kcal mol^{-1}.

State Si	ΔG_{Si-S0} (COSMO) (closed)	ΔG_{Si-S0} (COSMO) (open)	ΔG_{Si-S0} (DCOSMO-RS) (closed)	ΔG_{Si-S0} (DCOSMO-RS) (open)
S0	0.0	0.0	0.0	0.0
S1	29.3	33.1	31.4	33.4
S2	69.6	65.2	72.1	55.6
S3	103.9	118.6	103.9	114.3
S4	121.8	137.5	120.6	134.8

Table A.5: Relative free energies calculated for $\{Co_3Tm(pyMeO)_4\}$ employing either COSMO or DCOSMO-RS; Free energy differences are given in kcal mol^{-1}.

State Si	ΔG_{Si-S0} (COSMO) (closed)	ΔG_{Si-S0} (COSMO) (open)	ΔG_{Si-S0} (DCOSMO-RS) (closed)	ΔG_{Si-S0} (DCOSMO-RS) (open)
S0	0.0	0.0	0.0	0.0
S1	27.1	32.5	24.9	32.1
S2	62.5	80.8	56.6	68.6
S3	97.7	121.8	97.9	116.7
S4	119.7	136.1	115.6	132.7

A.2 Structure – Spin State

Table A.6: Energetically preferred spin state (number of (a)lpha and number of (b)beta electrons) on active center (Co1) per catalytic intermediate including explicit solvation (using CP2K). Note: The other transition metal and lanthanide centers were assumed to maintain a high spin configuration. The spin states reported here correspond to the initial guess used for the calculations, the resulting spin state after the wavefunction optimization can vary.

State Si	$\{Co_3Er\}$ (closed)	$\{Co_3Er\}$ (open)	$\{Co_3Tm(\}$ (closed)	$\{Co_3Tm\}$ (open)
S0	5a2b	5a2b	5a2b	5a2b
S1	5a1b	5a1b	4a2b	5a1b
S2	3a2b	3a2b	5a0b	3a2b
S3	3a3b	3a3b	5a1b	3a3b
S4	3a2b	3a2b	5a0b	3a2b

Table A.7: Energetically preferred spin state (number of (a)lpha and number of (b)beta electrons) on active center (Co1) per catalytic intermediate using explicit solvation structures where the water molecules were deleted (CP2K). Note: The other transition metal and lanthanide centers were assumed to maintain a high spin configuration. The spin states reported here correspond to the initial guess used for the calculations, the resulting spin state after the wavefunction optimization can vary.

State Si	$\{Co_3Er_4\}$ (closed)	$\{Co_3Er_4\}$ (open)	$\{Co_3Tm_4\}$ (closed)	$\{Co_3Tm_4\}$ (open)
S0	5a2b	5a2b	5a2b	5a2b
S1	5a1b	5a1b	4a2b	5a1b
S2	4a1b	5a0b	5a0b	4a1b
S3	5a1b	5a1b	5a1b	5a1b
S4	4a1b	5a0b	5a0b	5a0b

Table A.8: Energetically preferred spin state (number of (a)lpha and number of (b)beta electrons) on active center (Co1) per catalytic intermediate using implicit solvation (TURBOMOLE). Note: The other transition metal and lanthanide centers were assumed to maintain a high spin configuration. The spin states reported here correspond to the initial guess used for the calculations, the resulting spin state after the wavefunction optimization can vary.

State Si	$\{Co_3Er\}$ (closed)	$\{Co_3Er\}$ (open)	$\{Co_3Tm\}$ (closed)	$\{Co_3Tm\}$ (open)
S0	5a2b	5a2b	5a2b	5a2b
S1	4a2b	5a1b	3a3b	5a1b
S2	4a1b	3a2b	4a1b	4a1b
S3	3a3b	5a1b	5a1b	5a1b
S4	3a2b	5a0b	4a1b	5a0b

A.3 Structure – Selected Bond Lengths

Table A.9: Selected bond lengths of the open models of $\{Co_3Er(pyMeO)_4\}$ and $\{Co_3Tm(pyMeO)_4\}$ (given in Å).

State Si	$\{Co_3Er(pyMeO)_4\}$			$\{Co_3Tm(pyMeO)_4\}$		
spin state	Co1–O1	Co1–O2	Co1–O3	Co1–O1	Co1–O2	Co1–O3
S0-5a0b	3.44	1.98	3.37	3.41	1.98	3.39
S1-3a3b	3.42	1.96	3.56	3.42	1.96	3.42
S1-4a2b	3.41	1.91	4.06	3.40	1.95	3.47
S1-5a1b	3.41	1.98	3.55	3.40	1.97	3.48
S2-3a2b	3.53	2.10	3.29	3.53	2.09	3.28
S2-4a1b	3.35	2.01	3.57	3.50	2.01	3.57
S2-5a0b	3.49	2.03	3.51	3.49	2.02	3.54
S3-3a3b	3.33	1.95	3.57	3.30	1.93	3.51
S3-4a2b	3.31	1.95	3.60	3.34	1.95	3.40
S3-5a1b	3.29	1.94	3.37	3.30	1.94	3.40
S4-3a2b	3.34	1.94	3.48	3.28	1.98	3.64
S4-4a1b	3.39	1.96	3.44	3.35	1.96	3.51
S4-5a0b	3.33	1.94	3.38	3.30	1.95	3.40

Table A.10: Selected bond lengths of the closed models of $\{Co_3Er(pyMeO)_4\}$ and $\{Co_3Tm(pyMeO)_4\}$ (given in Å).

State Si	$\{Co_3Er(pyMeO)_4\}$			$\{Co_3Tm(pyMeO)_4\}$		
spin state	Co1–O1	Co1–O2	Co1–O3	Co1–O1	Co1–O2	Co1–O3
S0-5a0b	2.15	2.13	2.19	2.13	2.13	2.19
S1-3a3b	2.12	2.06	2.08	2.12	2.07	2.08
S1-4a2b	2.01	1.97	2.06	2.00	1.99	2.25
S1-5a1b	2.17	2.16	2.17	2.17	2.18	2.17
S2-3a2b	2.13	2.06	2.02	2.15	2.05	2.07
S2-4a1b	2.06	2.01	2.21	2.05	2.02	2.23
S2-5a0b	2.38	2.17	2.16	2.36	2.19	2.17
S3-3a3b	2.00	1.97	1.98	2.07	2.02	2.13
S3-4a2b	1.98	1.98	2.22	1.98	1.99	2.23
S3-5a1b	2.11	2.09	2.19	2.10	2.09	2.21
S4-3a2b	2.04	2.00	1.98	2.17	2.05	2.02
S4-4a1b	2.04	2.03	2.21	2.03	2.03	2.23
S4-5a0b	2.14	2.21	2.18	2.1	2.1	2.20

Table A.11: Shared electron numbers (SENs) for selected bonds within the cubane core of $\{Co_3Er(pyMeO)_4\}$ along the reaction path for the transition from the closed to the open structure.

SEN	closed	approx. transition state	open
$\sigma_{Co1-Co3}$	0.02	0.01	0.00
σ_{Co1-Er}	0.08	0.03	0.00
σ_{Co1-O1}	0.18	0.07	0.01
σ_{Co1-O2}	0.13	0.22	0.23
σ_{Co1-O3}	0.09	0.06	0.00

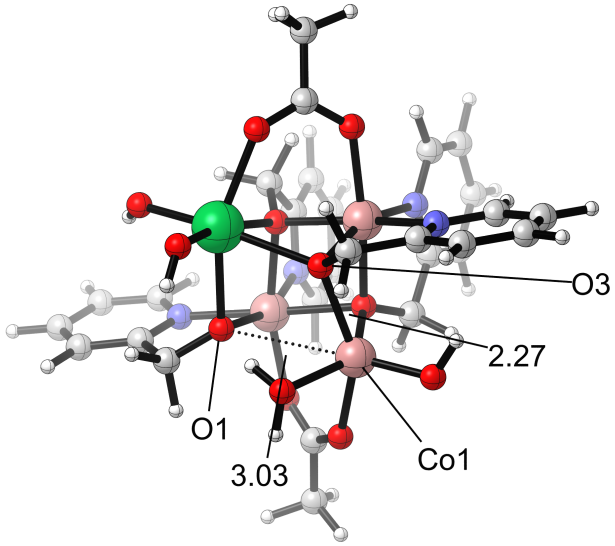


Figure A.1: Transition state for the opening of $\{Co_3Tm(pyMeO)_4\}$, obtained by optimizing the maximum electronic energy structure from the approximate reaction pathway. The saddle point character of the transition state was verified by a vibrational analysis, which gave one imaginary frequency, a combination of a Co1–O1 and a Co1–O3 stretching vibration. The electronic activation barrier (B3LYP-D3/COSMO) ΔE_{el}=3.9 kcal mol^{-1} is smaller than ΔE_{el}=6.5 kcal mol^{-1} obtained for the approximate transition state, thus supporting the conclusion that the barrier for the opening of the catalytic ground state is reasonably small.

A.4 Mixed Open/Closed Reaction Pathways

Table A.12: Mixed open/closed reaction pathways for $\{Co_3Er(pyMeO)_4\}$, the motifs are labeled as c(losed) or o(pen); Free energy differences are given in kcal mol^{-1}.

	S0	S1	S2	S3	S4
ΔG_{Si-S0}(CCCCC)	0.0	29.6	77.1	105.0	123.1
ΔG_{Si-S0}(**OO**CCC)	0.0	33.5	70.9	105.6	123.6
ΔG_{Si-S0}(**OOO**CC)	0.0	33.5	65.9	105.6	123.6
ΔG_{Si-S0}(C**OO**CC)	0.0	32.9	65.4	105.0	123.1
ΔG_{Si-S0}(CC**O**CC)	0.0	29.6	65.4	105.0	123.1
ΔG_{Si-S0}(**O**C**O**CC)	0.0	30.2	65.9	105.6	123.6
ΔG_{Si-S0}(**O**CCCC)	0.0	37.5	78.2	112.8	130.9

Table A.13: Mixed open/closed reaction pathways for $\{Co_3Tm(pyMeO)_4\}$, the motifs are labeled as c(losed) or o(pen); Free energy differences are given in kcal mol^{-1}.

	S0	S1	S2	S3	S4
ΔG_{Si-S0}(CCCCC)	0.0	27.4	63.2	98.7	121.0
ΔG_{Si-S0}(**OO**CCC)	0.0	28.1	76.9	98.7	121.0
ΔG_{Si-S0}(**OOO**CC)	0.0	23.8	81.6	103.5	125.7
ΔG_{Si-S0}(C**OO**CC)	0.0	32.8	67.9	103.5	125.7
ΔG_{Si-S0}(CC**O**CC)	0.0	32.2	67.9	103.5	125.7
ΔG_{Si-S0}(**O**C**O**CC)	0.0	23.3	76.9	98.7	121.0
ΔG_{Si-S0}(**O**CCCC)	0.0	32.2	81.6	103.5	125.7

A.5 Structures – Open and Closed Cage Structures

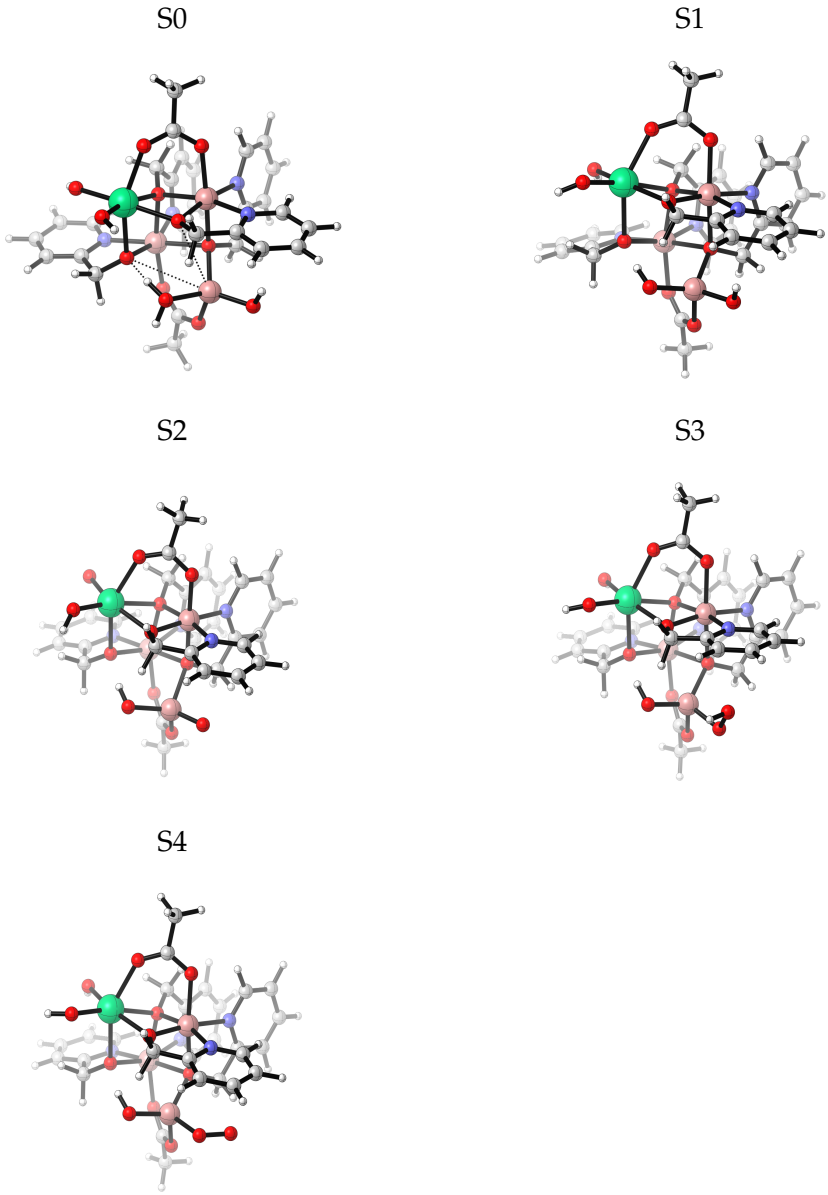


Figure A.2: Open model of $\{Co_3Er(pyMeO)_4\}$ optimized with TURBOMOLE using the implicit solvent continuum model COSMO.

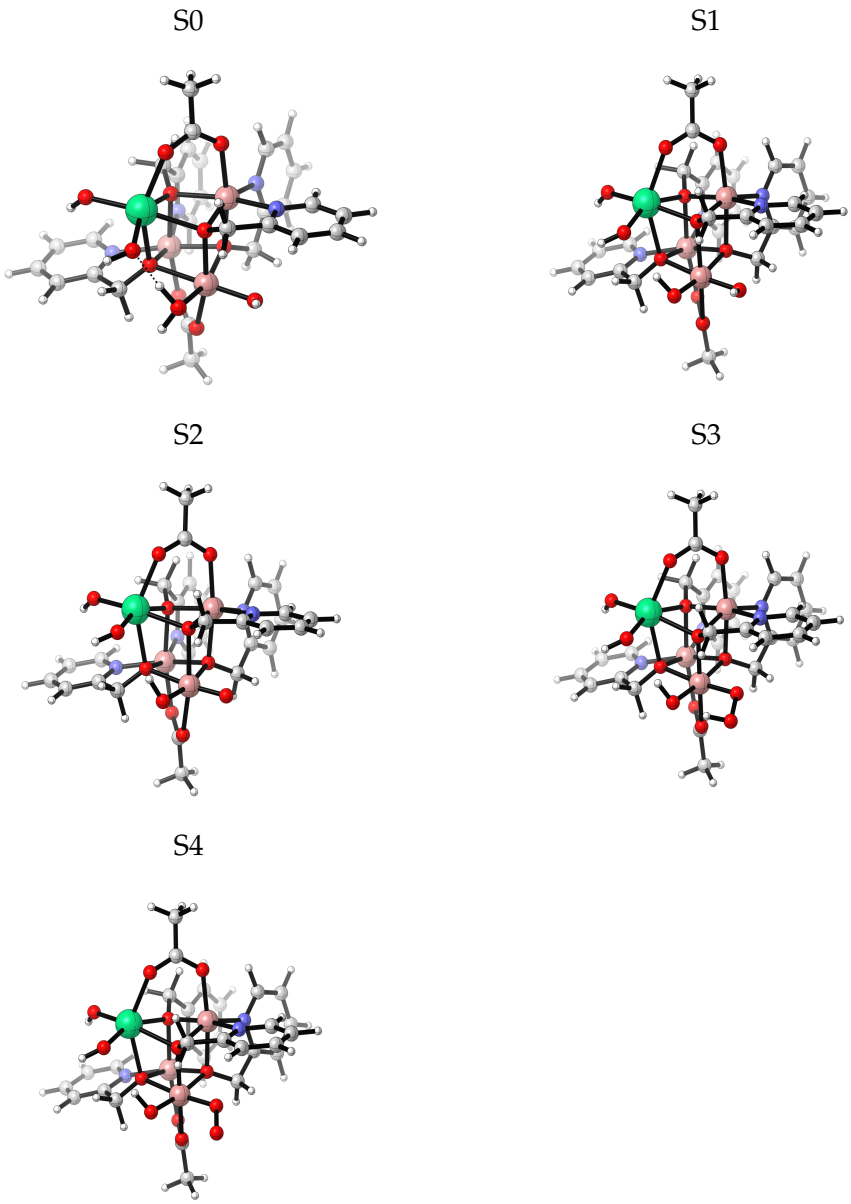
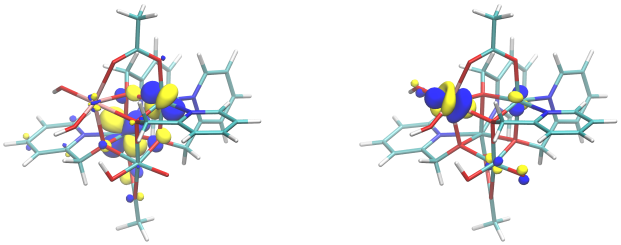


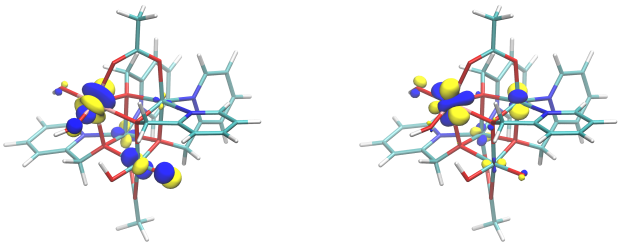
Figure A.3: Closed model of $\{Co_3Er(pyMeO)_4\}$ optimized with TURBOMOLE using the implicit solvent continuum model COSMO.

A.6 Frontier Orbitals – Intermediate S2

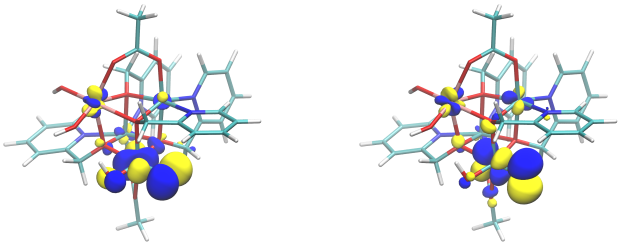
For each structure the HOMO as well as the first five LUMOs are shown. Since the orbitals originate from unrestricted Kohn-Sham DFT each of them is associated with a spin a(lpha) or b(eta).



HOMO(b) LUMO(b)

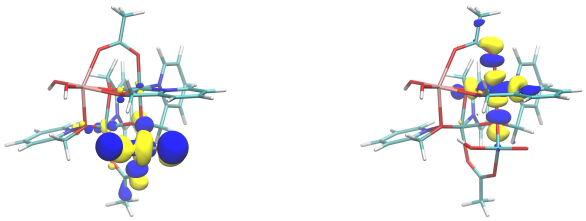


LUMO+1(b) LUMO+2(b)

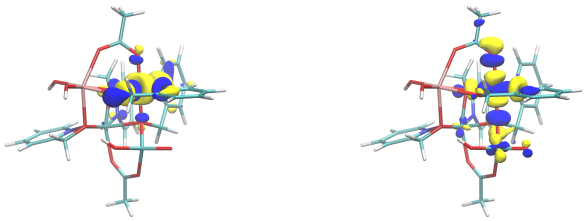


LUMO+3(b) LUMO+4(b)

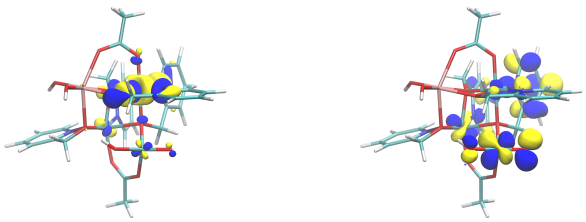
Figure A.4: Canonical molecular orbitals of the closed structure of $\{Co_3Er(pyMeO)_4\}$ in the S2 state optimized with TURBOMOLE using the implicit solvent continuum model COSMO (isosurface 0.035 $e^-/Bohr^3$(blue), -0.035 $e^-/Bohr^3$ (yellow)). The left column represents the alpha channel (a) and the right the beta channel (b) of the unrestricted Kohn-Sham DFT calculations.



HOMO(a) LUMO(a)

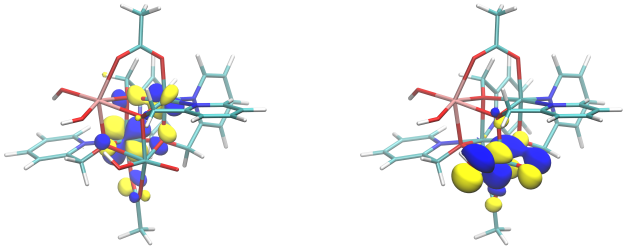


LUMO+1(a) LUMO+1(b)

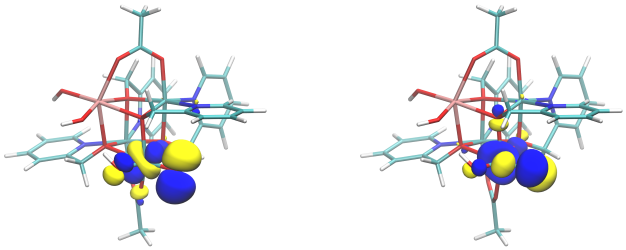


LUMO+2(b) LUMO+3(a)

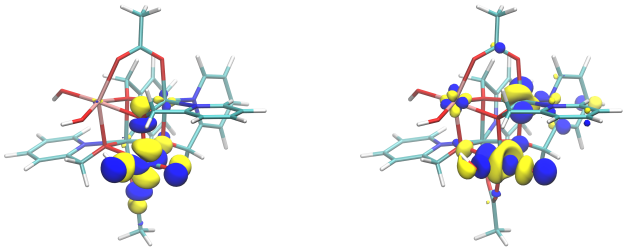
Figure A.5: Canonical molecular orbitals of the closed structure of $\{Co_3Er(pyMeO)_4\}$ in the S2 state optimized with TURBOMOLE using the implicit solvent continuum model COSMO (isosurface 0.035 e^-Bohr^3(blue), -0.035 $e^-/Bohr^3$ (yellow)). The left column represents the alpha channel (a) and the right the beta channel (b) of the unrestricted Kohn-Sham DFT calculations.



HOMO(b) LUMO(b)

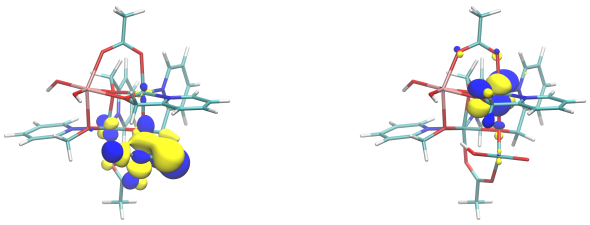


LUMO+1(b) LUMO+2(b)

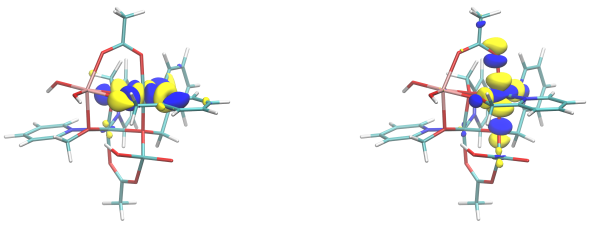


LUMO+3(b) LUMO+4(b)

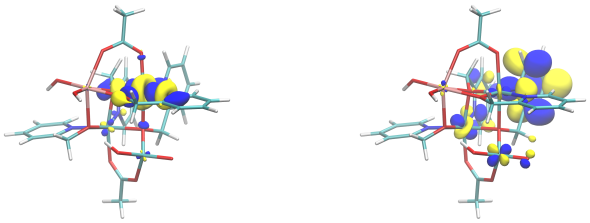
Figure A.6: Canonical molecular orbitals of the closed structure of $\{Co_3Tm(pyMeO)_4\}$ in the S2 state optimized with TURBOMOLE using the implicit solvent continuum model COSMO (isosurface 0.035 $e^-/Bohr^3$(blue), -0.035 $e^-/Bohr^3$ (yellow)). The left column represents the alpha channel (a) and the right the beta channel (b) of the unrestricted Kohn-Sham DFT calculations.



HOMO(b) LUMO(b)



LUMO(a) LUMO+1(b)



LUMO+2(b) LUMO+3(b)

Figure A.7: Canonical molecular orbitals of the closed structure of $\{Co_3Tm(pyMeO)_4\}$ in the S2 state optimized with TURBOMOLE using the implicit solvent continuum model COSMO (isosurface 0.035 $e^{-}Bohr^3$(blue), -0.035 $e^{-}/Bohr^3$ (yellow)). The left column represents the alpha channel (a) and the right the beta channel (b) of the unrestricted Kohn-Sham DFT calculations.

A.7 Structural Analysis of the Models with Explicit Solvation

A detailed description of the simulations employing explicit solvation can be found in the Supplementary Information of Schilling, M.; Hodel, F. H.; and Luber, S.; *ChemSusChem* **2017** *10*, 4561–4569.[133].

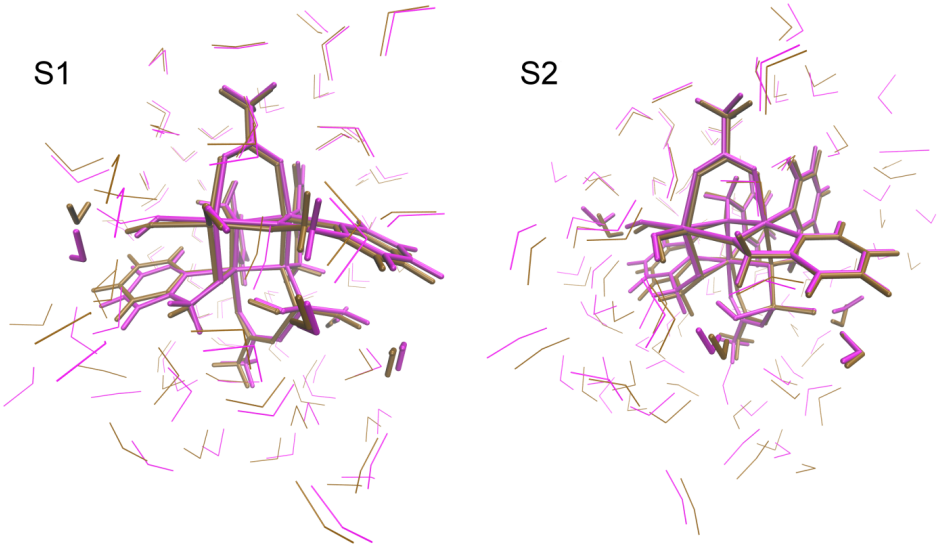


Figure A.8: Comparison of the open structures of $\{Co_3Er(pyMeO)_4\}$ (magenta) and $\{Co_3Tm(pyMeO)_4\}$ (ochre) in S1 and S2.

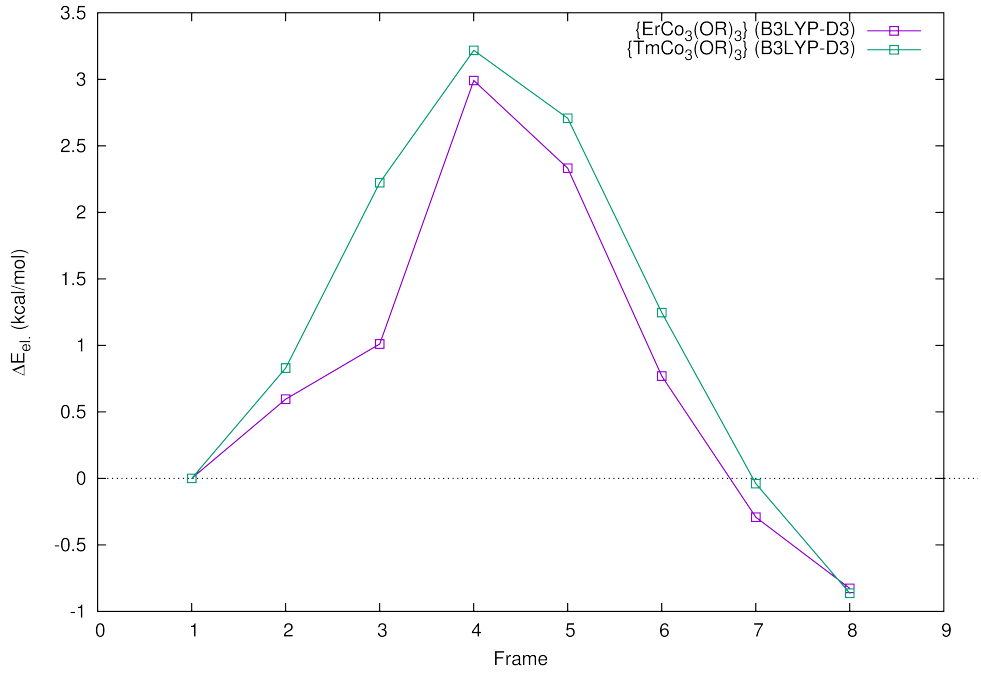


Figure A.9: Optimized reaction path for the opening of the catalytic ground state of $\{Co_3Ln(pyMeO)_4\}$ (where Ln = Tm and Er) using NEB calculations.

B Rational Design of the Py5–Ligand Framework for Ru–Based WOCs

B.1 Oxygen in cis Position to Chloride

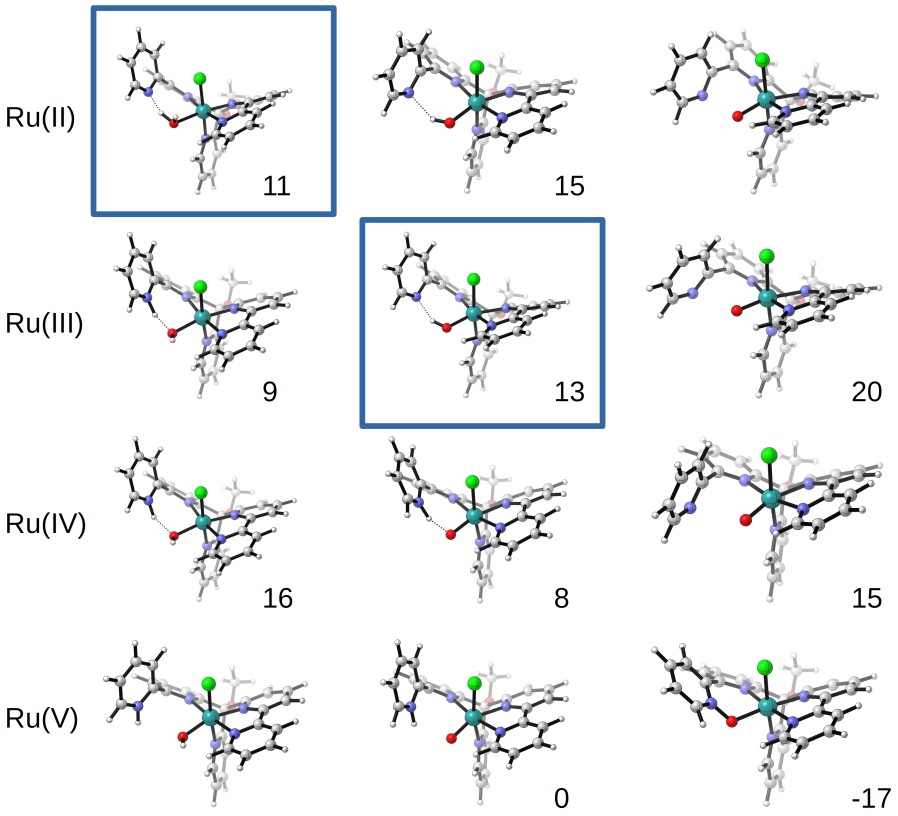


Figure B.1: Isomers of complexes with the Py5OMe ligand where the entering water molecule is in *cis* position to the chloride. Energies are given in kcal mol^{-1}, relative to the structures with the oxygen *trans* to chloride (B3LYP-D3/def2-TZVP/COSMO//BP86-D3/def2-TZVP). Structures are sequentially deprotonated from left to right. It is noticeable that in these arrangements the proton is more likely to shift from the oxygen to the pyridyl fragment. Highlighted with blue rectangles are two of the structures of the thermodynamic cycle which are comparable for *cis* and *trans* in terms of oxidation state and protonation pattern. The energies of the *cis* isomers are higher than those for the *trans* isomers, therefore it is unlikely that the catalytic cycle involves any of the above species. Notice how in the $Ru^{V}-O$ structure the ligand strain has led to a nucleophilic attack on the oxygen. This might be an artifact of the absence of explicit solvent molecules.

B.2 Oxygen Evolution

Because of the large dependence on explicit solvation, structures with one more additional solvent molecule for the TS of Py5OMe were prepared and optimized (3 water molecules = 1 for WNA + 2 for solvation). We limited ourselves to structures where the second explicit water molecule belongs to the first solvation shell as seen

in Figure B.2 in the manuscript. The activation energies for these structures are quite similar to those obtained with only one explicit solvent molecule (15.6 kcal mol^{-1}) and are in the range of 14.3 to 17.1 kcal mol^{-1} with respect to the separated reactants, or between 5 to 10 kcal mol^{-1} with respect to the associated complex (table B.2). Regarding the value with respect to the separated reactants, it appears that there is no need for a second explicit water molecule as the activation energies are quite similar. However, with static calculations, it is in general difficult to meaningfully sample all possible starting guesses. The structures with two explicit solvent molecules resemble those with only one very much, because we used them as starting guesses. Therefore, with different starting guesses one might obtain other structures with a significantly different pattern which may lead to other values for the activation energies. For such a task molecular dynamics or other techniques may be more suitable.

Ligand	Ru^{II}-Py	Ru^{III}-Py	Ru^{II}-Cl	Ru^{III}-Cl
Py5Me	14.2	11.5	7.6	19.5
Py5OMe	4.4	1.4	8.4	21.6
Py5Et	7.3	3.6	7.8	20.3

Table B.1: Energies (kcal mol^{-1}) of chlorido-water exchange (Ru-Cl) and the pyridine-water exchange (Ru-py) with the different ligands and the complexes with ruthenium in oxidation state II or III (values for Py5OMe and Py5Me were previously reported by Gil-Sepulcre *et al.*[173]).

Variant	AR	TS	AP
A	7.1	14.3	7.1
B	10.5	15.1	7.8
C	10.9	17.1	8.2

Table B.2: Energies (kcal mol^{-1}) of WNA structures with Py5OMe containing two explicit water molecules. Energies are given relative to the free $OMe-N^4-Ru^V-O-Cl$ complex and three free water molecules. AR = associated reactants, TS = transition state, AP = associated products.

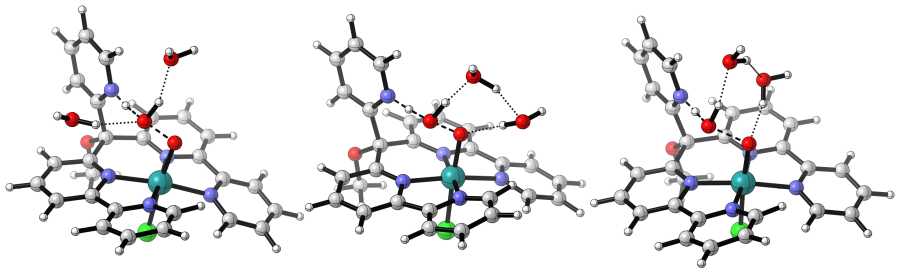


Figure B.2: Transition-states for OMe–N^4–Ru^V–O–Cl with two explicit water molecules in the first solvation shell. From left to right: variant A, B and C.

B.3 Dissociation Scans

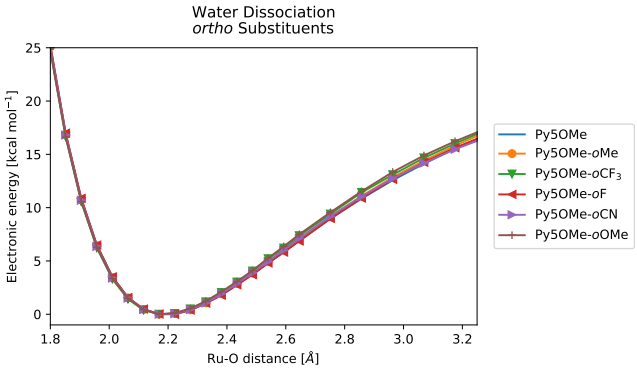


Figure B.3: Water dissociation scans with the indicated substituents in the *ortho* position of the axial pyridine.

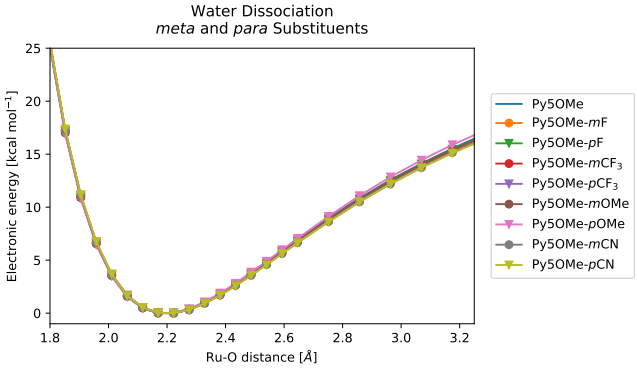


Figure B.4: Water dissociation scans with the indicated substituents in *meta* and *para* positions of the axial pyridine.

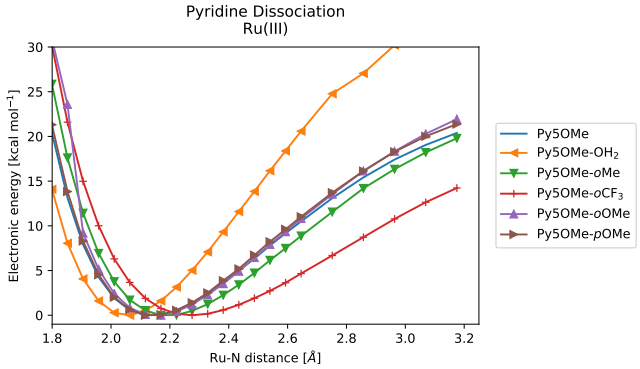


Figure B.5: Pyridine dissociation scans on Ru^{III}. Note the preserved energy ordering (as in Ru^{II}) *ortho*-CF_3 ¡ *ortho*-Me ¡ Py5OMe ≈ *ortho*-OMe ≈ *para*-OMe ¡ Py5Me. The *trans*-aqua species are also included.

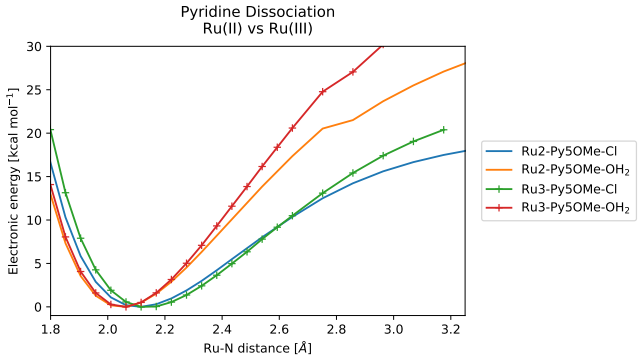


Figure B.6: Pyridine dissociations scans comparing Ruthenium in oxidations states II and III. OH_2 and Cl stand for the ligand in *trans* position.

B.4 WNA Correlation

WNA correlation

Mod.	AR	TS	AP	Ru-spin (x10)	O-spin (x10)	LUMO (x20)	d(N-O)
o-OMe	7.30	14.84	7.69	5.25	4.46	-3.95	2.63
m-OMe	6.46	14.48	6.14	6.26	3.79	-3.81	2.41
p-OMe	6.17	11.71	3.22	7.58	2.84	-3.72	2.36
o-F	7.18	17.10	15.50	4.98	4.68	-4.04	2.71
m-F	6.78	14.67	10.35	5.72	4.06	-3.91	2.48
p-F	6.50	13.45	7.68	6.09	3.76	-3.87	2.44
o-CF_3	9.23	20.86	17.11	5.02	4.66	-4.03	2.69
m-CF_3	7.33	16.35	11.61	5.63	4.12	-3.92	2.47
p-CF_3	6.54	14.79	10.32	5.83	3.95	-3.91	2.46

Table B.3: Energies for AR, TS and AP in kcal mol^{-1}; Mulliken spin-densities for Ru^{V}=O complex (scaled by 10), LUMO energy (eV, scaled by 20); distance of pyridine-N to oxo in Å.

C Determination of pK_a Values via ab initio Molecular Dynamics

C.1 Potentials of Mean Force

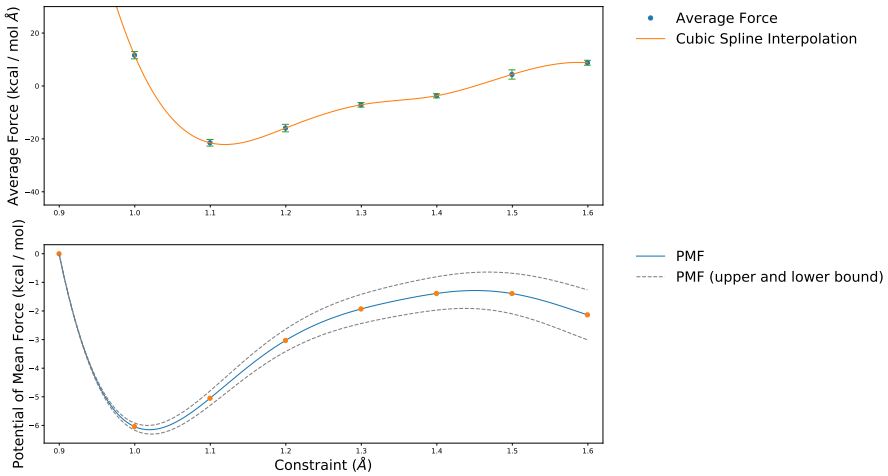


Figure C.1: Average force and PMF profile of HCOOH in a cubic box with a side length of 15.6406 Å.

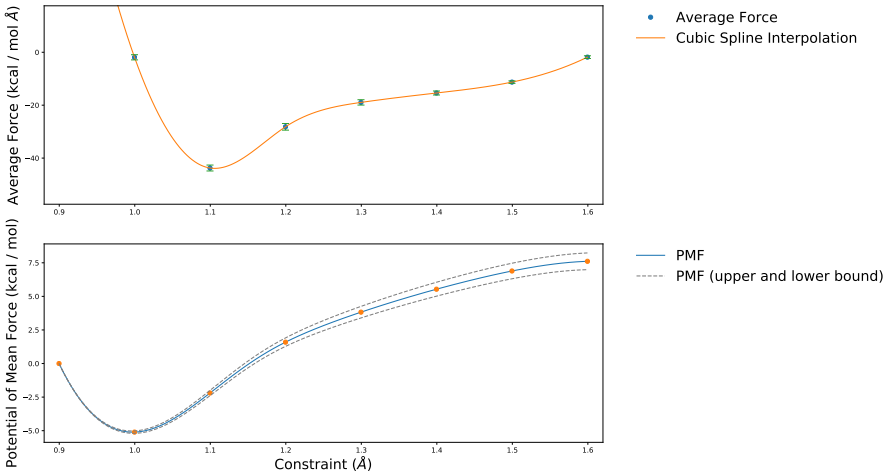


Figure C.2: Average force and PMF profile of PhOH in a cubic box with a side length of 15.6406 Å.

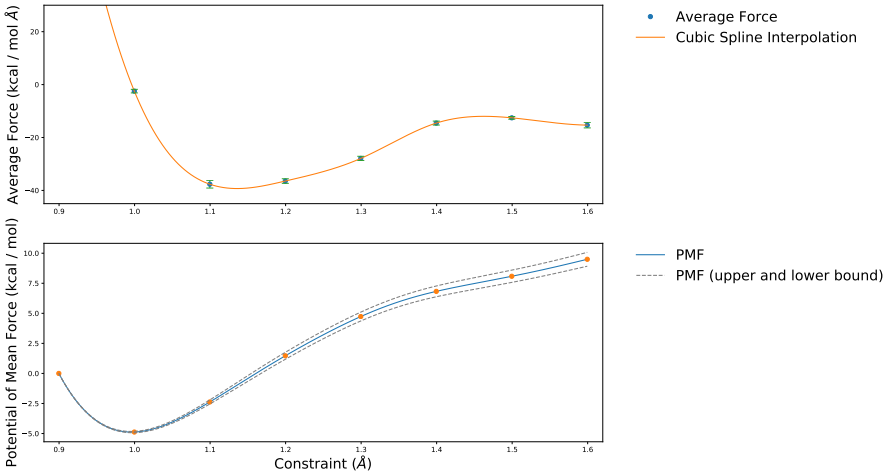


Figure C.3: Average force and PMF profile of $[Ru^{II}Py_5Me(H_2O)]^{2+}$ in a cubic box with a side length of 15.6406 Å.

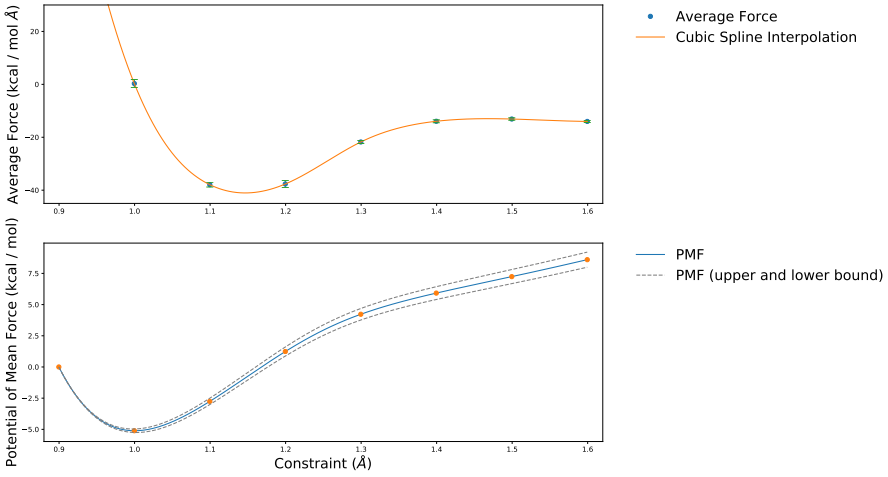


Figure C.4: Average force and PMF profile of $[Ru^{II}Py_5OMe(H_2O)]^{2+}$ in a cubic box with a side length of 15.6406 Å.

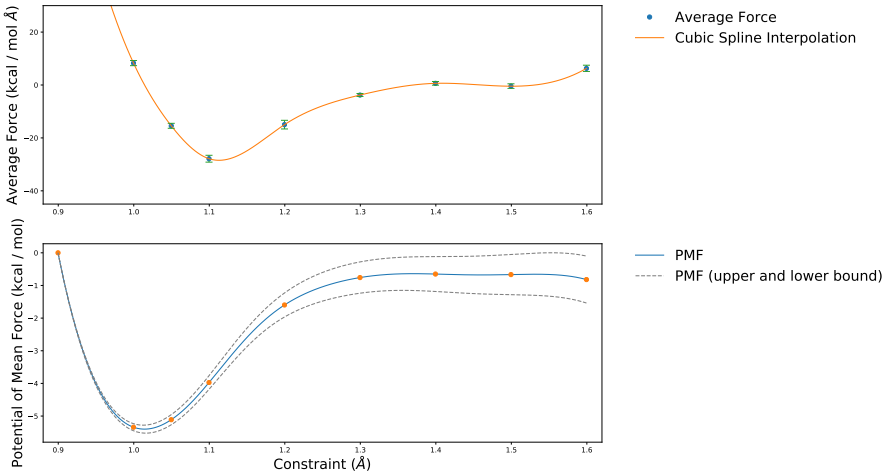


Figure C.5: Average force and PMF profile of $[Ru^{III}Py_5OMe(H_2O)]^{3+}$ in a cubic box with a side length of 15.6406 Å.

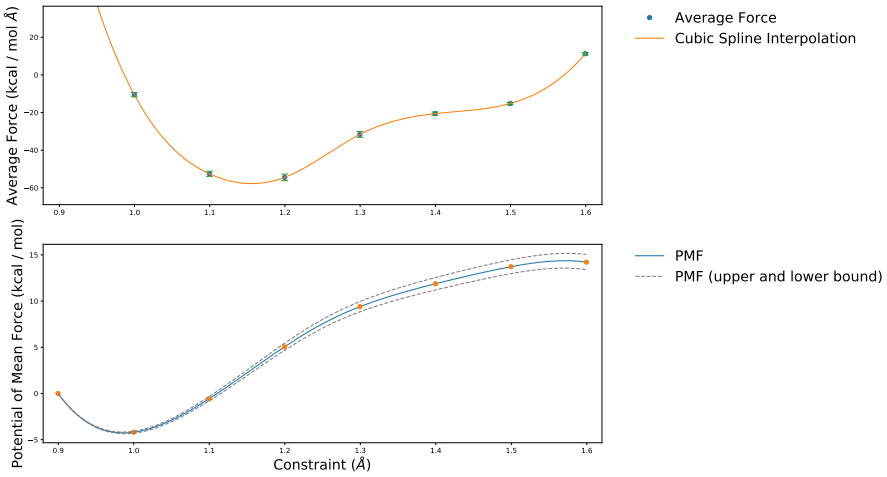


Figure C.6: Average force and PMF profile of water in a cubic box with a side length of 19.7340 Å.

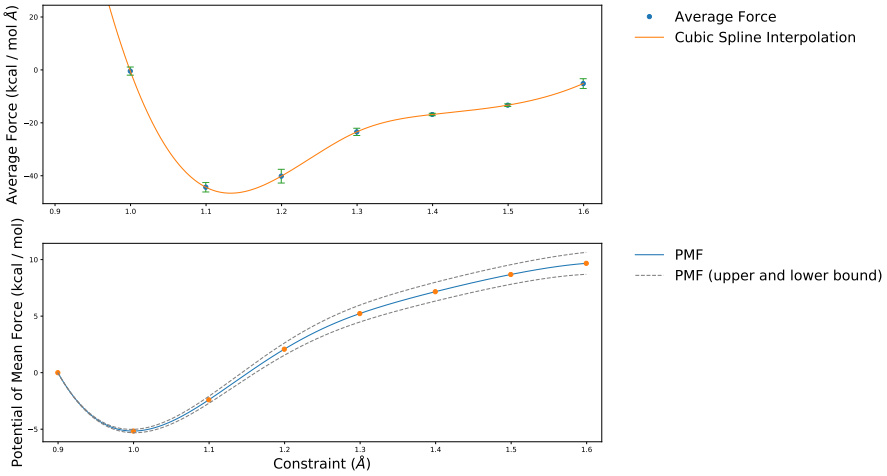


Figure C.7: Average force and PMF profile of $[Ru^{II}Py_5OMe(H_2O)]^{2+}$ in a cubic box with a side length of 19.7340 Å.

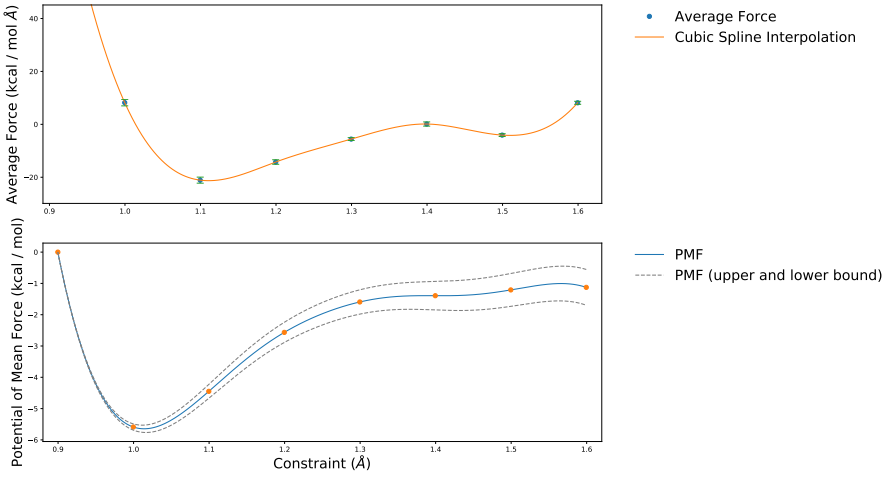


Figure C.8: Average force and PMF profile of $[Ru^{III}Py_5OMe(H_2O)]^{3+}$ in a cubic box with a side length of 19.7340 Å.

C.2 Absolute and Probabilistic pKa

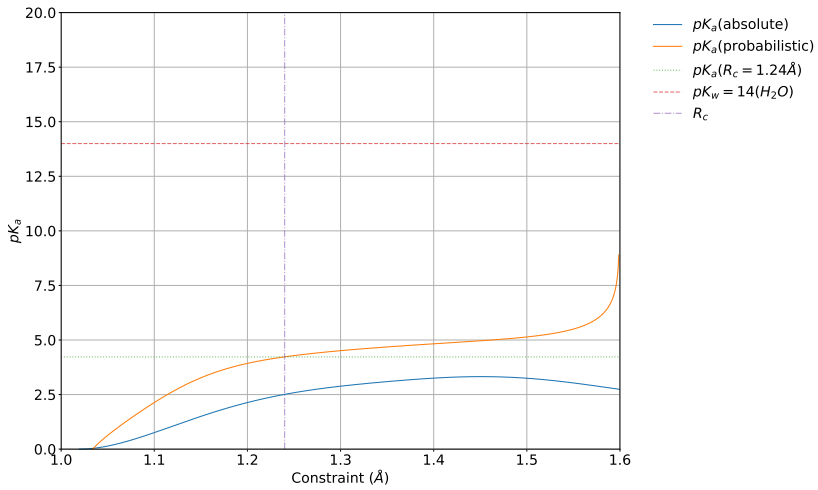


Figure C.9: Absolute and probabilistic pK_a calculated for HCOOH in a cubic box with a side length of 15.6406 Å.

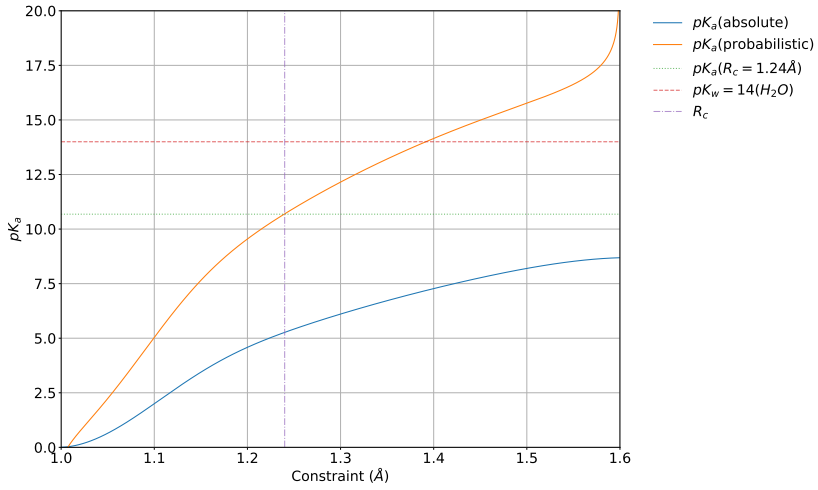


Figure C.10: Absolute and probabilistic pK_a calculated for PhOH in a cubic box with a side length of 15.6406 Å.

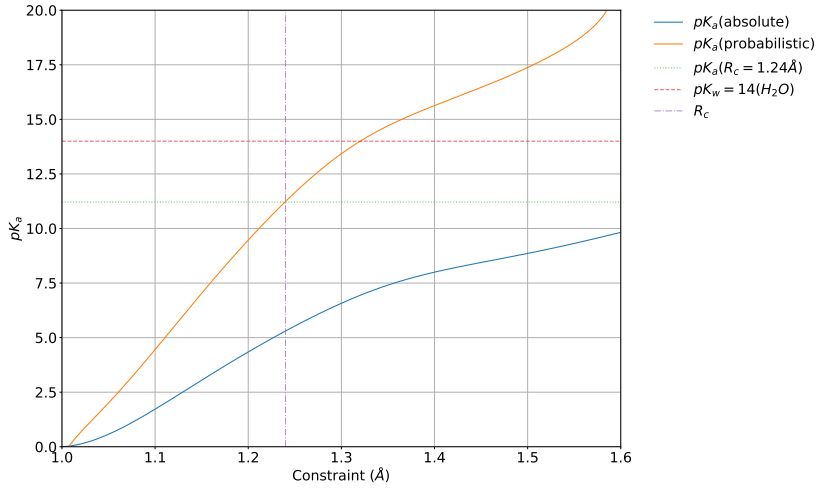


Figure C.11: Absolute and probabilistic pK_a calculated for $[Ru^{II}Py_5Me(H_2O)]^{2+}$ in a cubic box with a side length of 15.6406 Å.

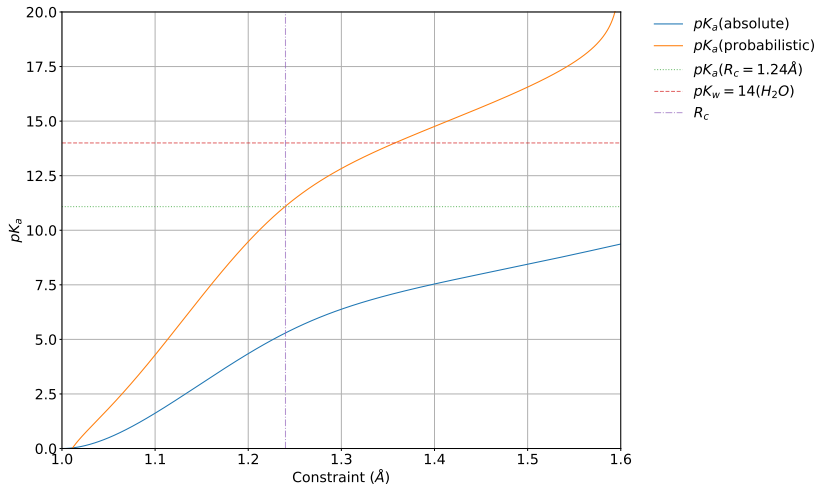


Figure C.12: Absolute and probabilistic pK_a calculated for $[Ru^{II}Py_5OMe(H_2O)]^{2+}$ in a cubic box with a side length of 15.6406 Å.

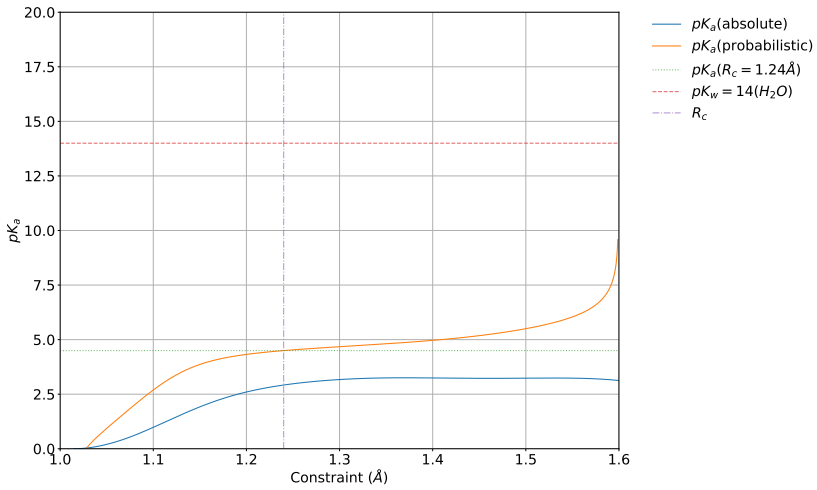


Figure C.13: Absolute and probabilistic pK_a calculated for $[Ru^{III}Py_5OMe(H_2O)]^{3+}$ in a cubic box with a side length of 15.6406 Å.

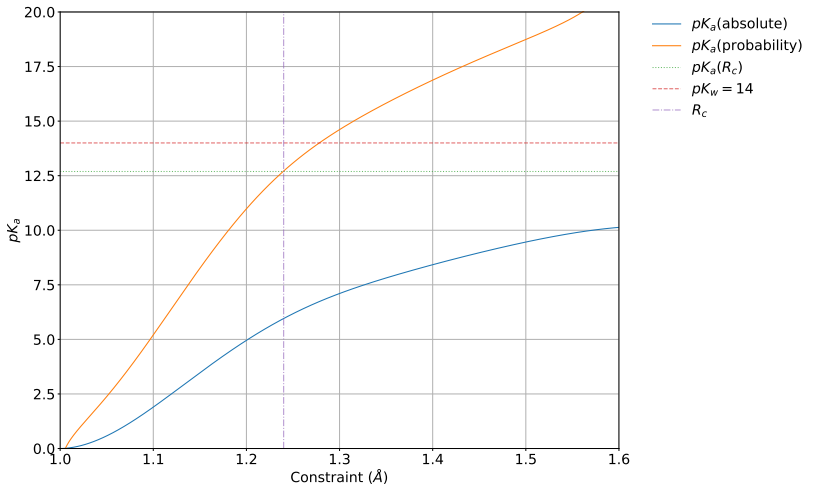


Figure C.14: Absolute and probabilistic pK_a calculated for $[Ru^{II}Py_5OMe(H_2O)]^{2+}$ in a cubic box with a side length of 19.7340 Å.

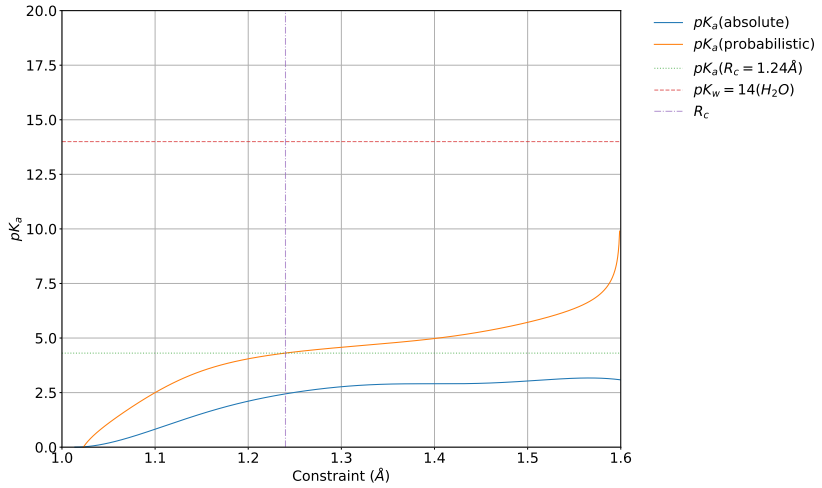


Figure C.15: Absolute and probabilistic pK_a calculated for $[Ru^{III}Py_5OMe(H_2O)]^{3+}$ in a cubic box with a side length of 19.7340 Å.

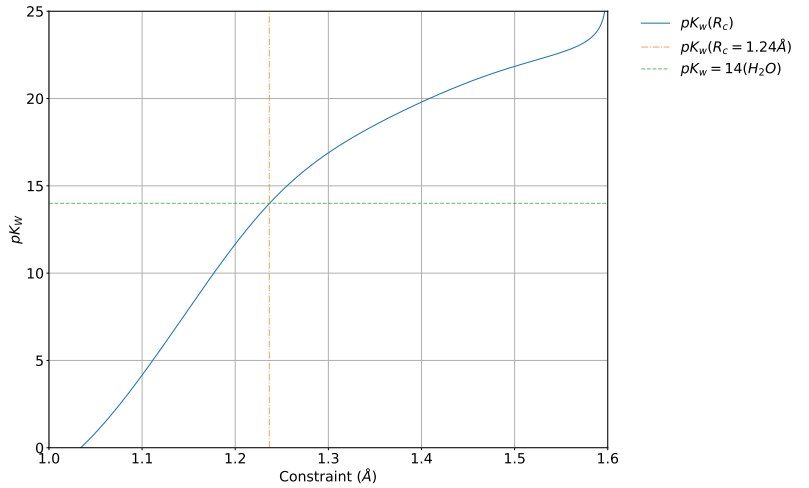


Figure C.16: Determination of R_c from the simulation of water, by fitting to the experimental value in box with a side length of 19.7340 Å.

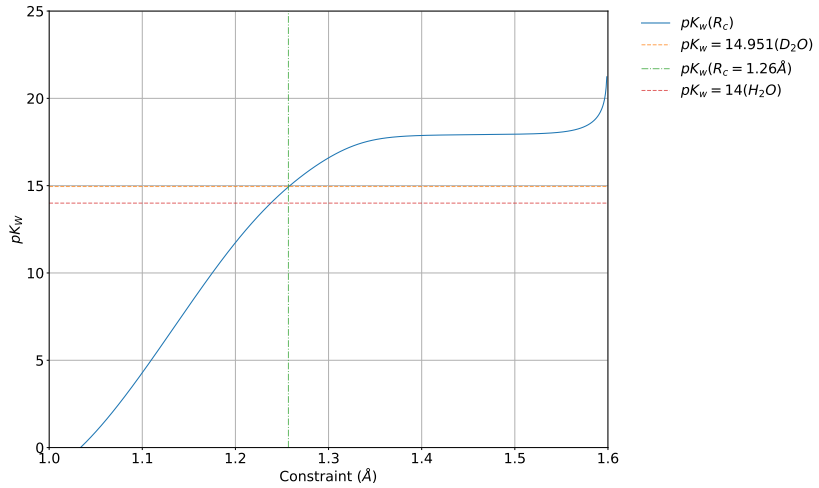


Figure C.17: Determination of R_c from the simulation of water, by fitting to the experimental value of D_2O in box with a side length of 15.6406 Å.

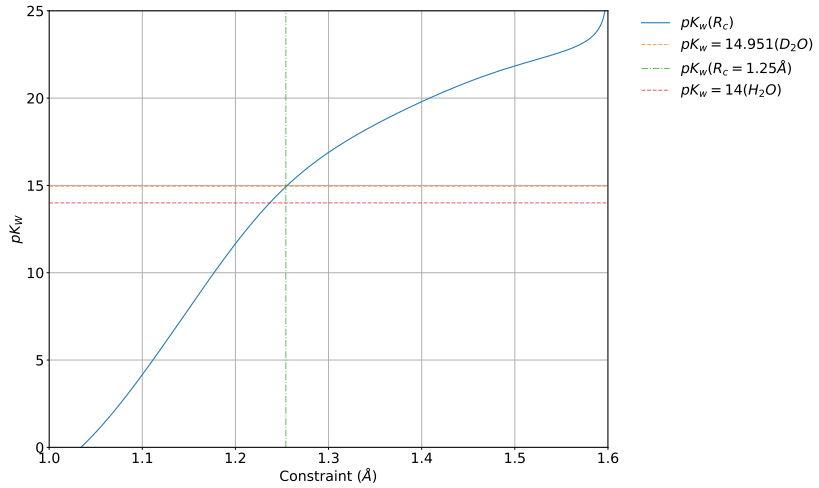


Figure C.18: Determination of R_c from the simulation of water, by fitting to the experimental value of D_2O in box with a side length of 19.7340 Å.

C.3 Probabilistic pKa - Dependence on R_c

Table C.1: All results presented here are calculated at 320 K, for a cubic box with a side length of 15.6406 Å, over a trajectory of about 20 ps (where the first 5 ps were not included in the evaluation). The pK_a values were calculated using the probabilistic method for different values of R_c. The '*' indicates the that R_c value determined was determined from clean water using the same settings.

Molecule	1.19 Å	1.22 Å	1.24* Å	1.28 Å	1.30 Å
H_2O	-	-	-	-	-
HCOOH	3.8	4.1	4.2	4.4	4.5
PhOH	9.2	10.1	10.7	11.7	12.1
$[Ru^{II}Py_5Me(H_2O)]^{2+}$	9.0	10.4	11.2	12.8	13.4
$[Ru^{II}Py_5OMe(H_2O)]^{2+}$	9.0	10.3	11.1	12.3	12.8
$[Ru^{III}Py_5OMe(H_2O)]^{3+}$	4.3	4.4	4.5	4.6	4.7

Table C.2: All results presented here are calculated at 320 K, for a cubic box with a side length of 19.7340 Å, over a trajectory of about 20 ps (where the first 5 ps were not included in the evaluation). The pK_a values were calculated using the probabilistic method for different values of R_c. The '*' indicates the that R_c value determined was determined from clean water using the same settings.

Molecule	1.22 Å	1.24* Å	1.28 Å	1.30 Å
H_2O	-	-	-	-
$[Ru^{II}Py_5OMe(H_2O)]^{2+}$	11.9	12.7	14.0	14.6
$[Ru^{III}Py_5OMe(H_2O)]^{3+}$	4.2	4.3	4.5	4.6

C.4 Absolute pKa - Convergence

The value for the absolute pK_a converges quickly to a constant value, however there is no guarantee that the dissociated state (i.e. A−H = 1.6 Å) is always the same (reprotonation, proton hopping).

Table C.3: Absolute pK_a calculated at 320 K, for a cubic box with a side length of 15.6406 Å and a cutoff a cutoff R_c = 1.24 Å, increasing the equilibration time (ps) - total length of all trajectories is about 20 ps minus the equilibration time.

equilibration time	2.5 ps	5 ps	7.5 ps	10 ps	12.5 ps
HCOOH	2.7	2.7	2.7	2.3	1.9
PhOH	8.8	8.7	8.7	8.8	8.6
$[Ru^{II}Py_5Me(H_2O)]^{2+}$	9.7	9.8	9.9	10.0	10.1
$[Ru^{II}Py_5OMe(H_2O)]^{2+}$	9.4	9.4	9.4	9.4	9.6
$[Ru^{III}Py_5OMe(H_2O)]^{3+}$	3.0	3.1	3.3	3.5	3.4

C.5 Probabilistic pKa - Convergence

The value for the probabilistic pK_a requires more steps to converge, in particular for the highly charged system.

Table C.4: probabilistic pK_a calculated at 320 K, for a cubic box with a side length of 15.6406 Å and a cutoff R_c = 1.24 Å, increasing the equilibration time (ps) - total length of all trajectories is about 20 ps minus the equilibration time.

equilibration time	2.5 ps	5 ps	7.5 ps	10 ps	12.5 ps
HCOOH	4.1	4.2	4.2	3.9	3.4
PhOH	11.0	10.7	10.6	10.7	10.5
$[Ru^{II}Py_5Me(H_2O)]^{2+}$	11.0	11.2	11.5	11.7	11.6
$[Ru^{II}Py_5OMe(H_2O)]^{2+}$	11.1	11.1	11.3	11.4	11.5
$[Ru^{III}Py_5OMe(H_2O)]^{3+}$	4.4	4.5	4.8	5.1	4.7

C.6 Relative pKa - Convergence

Table C.5: Relative pK_a calculated at 320 K, for a cubic box with a side length of 15.6406 Å and a cutoff R_c = 1.24 Å, increasing the equilibration time (ps) - total length of all trajectories is about 20 ps minus the equilibration time.

equilibration time	2.5 ps	5 ps	7.5 ps	10 ps	12.5 ps
HCOOH	6.7	6.8	6.6	6.4	6.0
PhOH	12.6	12.6	12.6	12.8	12.5
$[Ru^{II}Py_5Me(H_2O)]^{2+}$	13.5	13.7	13.8	13.9	14.0
$[Ru^{II}Py_5OMe(H_2O)]^{2+}$	13.3	13.3	13.3	13.4	13.5
$[Ru^{III}Py_5OMe(H_2O)]^{3+}$	6.9	7.1	7.3	7.5	7.4

C.7 Relative pKa - Reduced Constraint

Table C.6: Relative pK_a calculated at 320 K, for a cubic box with a side length of 15.6406 Å and a cutoff R_c = 1.24 Å, where only constraints within 0.9 to 1.4 Å where considered for the water reference.

molecule	pK_a
HCOOH	5.8
PhOH	11.8
$[Ru^{II}Py_5Me(H_2O)]^{2+}$	12.0
$[Ru^{II}Py_5OMe(H_2O)]^{2+}$	12.3
$[Ru^{III}Py_5OMe(H_2O)]^{3+}$	6.2

C.8 Implications of using D_2O

Absolute pKa

If the pK_a values are considered to be p$K_a{}^D$ values due to the use of molecular mass of deuterium, then the absolute pK_a values have to be corrected according to the linear relation discussed beforehand. Applying those conversion factors lowers all pK_a values and leads in most cases to a worse agreement with the experiment.

Table C.7: All results presented here are calculated at 320 K, for a cubic box with side length of 15.6406 Å, over a trajectory of about 20 ps (where the first 5 ps were not included in the evaluation). The p$K_a{}^D$ obtained using the absolute method were converted to p$K_a{}^H$ values according to 6.10[a] and Equation 6.13 [b].

Molecule	pK_a (exp.)	p$K_a{}^D$	p$K_a{}^H[a]$	p$K_a{}^H[b]$
H_2O	14.0	-	-	-
HCOOH	3.8[317]	2.7	2.3	2.6
PhOH	10.0[317]	8.7	8.0	8.1
$[Ru^{II}Py_5Me(H_2O)]^{2+}$	∼11[172]	9.8	9.1	9.2
$[Ru^{II}Py_5OMe(H_2O)]^{2+}$	∼11[172]	9.3	8.6	8.7
$[Ru^{III}Py_5OMe(H_2O)]^{3+}$	∼2.5[172]	3.1	2.7	2.9

Table C.8: All results presented here are calculated at 320 K, for a cubic box with side length of 19.7340 Å, over a trajectory of about 20 ps (where the first 5 ps were not included in the evaluation). The p$K_a{}^D$ obtained using the absolute method were converted to p$K_a{}^H$ values according to 6.10[a] and Equation 6.13 [b].

Molecule	pK_a (exp.)	p$K_a{}^D$	p$K_a{}^H[a]$	p$K_a{}^H[b]$
H_2O	14.0	-	-	-
$[Ru^{II}Py_5OMe(H_2O)]^{2+}$	∼11[172]	10.1	9.4	9.4
$[Ru^{III}Py_5OMe(H_2O)]^{3+}$	∼2.5[172]	3.1	2.7	2.9

Relative pK_a

As for the probabilistic pK_a values the pK_a of D_2O had to be considered for the calculation of the relative pK_a values. The comparison of p$K_a{}^H$ values obtained by converting p$K_a{}^D$ values with p$K_a{}^H$ obtained by referencing to H_2O revealed an overall reduction of the pK_a values by more than 1 unit. As a consequence there is a better agreement for basic compounds.

Table C.9: All results presented here are calculated at 320 K, for a cubic box with side length of 15.6406 Å, over a trajectory of about 20 ps (where the first 5 ps were not included in the evaluation). All pK_a values were obtained using the relative method. The p$K_a{}^H[a]$ values were calculated relative to H_2O, while p$K_a{}^D$ were calculated relative to D_2O. The latter were converted into p$K_a{}^H$ values according to 6.10[a] and Equation 6.13 [b].

Molecule	pK_a (exp.)	p$K_a{}^H[a]$	p$K_a{}^D$	p$K_a{}^H[b]$	p$K_a{}^H[c]$
H_2O	14.0	-	-	-	-
HCOOH	3.8[317]	6.7	5.7	5.2	5.5
PhOH	10.0[317]	12.5	11.5	10.8	10.8
$[Ru^{II}Py_5Me(H_2O)]^{2+}$	∼11[172]	13.7	12.7	11.9	11.9
$[Ru^{II}Py_5OMe(H_2O)]^{2+}$	∼11[172]	13.3	12.3	11.5	11.5
$[Ru^{III}Py_5OMe(H_2O)]^{3+}$	∼2.5[172]	7.1	6.1	5.6	5.8

Table C.10: All results presented here are calculated at 320 K, for a cubic box with side length of 19.7340 Å, over a trajectory of about 20 ps (where the first 5 ps were not included in the evaluation). All pK_a values were obtained using the relative method. The p$K_a{}^H[a]$ values were calculated relative to H_2O, while p$K_a{}^D$ were calculated relative to D_2O. The latter were converted into p$K_a{}^H$ values according to 6.10[a] and Equation 6.13 [b].

Molecule	pK_a (exp.)	p$K_a{}^H[a]$	p$K_a{}^D$	p$K_a{}^H[b]$	p$K_a{}^H[c]$
H_2O	14.0	-	-	-	-
$[Ru^{II}Py_5OMe(H_2O)]^{2+}$	∼11[172]	11.5	10.6	9.9	9.9
$[Ru^{III}Py_5OMe(H_2O)]^{3+}$	∼2.5[172]	4.6	3.7	3.3	3.5

C.9 Overview Simulation Time

Table C.11: Number of simulation steps per model system / constraint for a cubic box with a side length of 15.6406 Å. Each time step corresponds to 0.5 fs. Abbreviations: $Ru^{II}Me$ = $[Ru^{II}Py_5Me(H_2O)]^{2+}$, $Ru^{II}OMe$ = $[Ru^{II}Py_5OMe(H_2O)]^{2+}$, $Ru^{III}OMe$ = $[Ru^{III}Py_5OMe(H_2O)]^{3+}$

d(A−H)	H_2O	PhOH	HCOOH	$Ru^{II}Me$	$Ru^{II}OMe$	$Ru^{III}OMe$
0.9Å	30529	38380	39245	39731	40735	43373
1.0Å	45516	37244	40913	38672	40224	43360
1.1Å	44861	40056	40233	48999	40268	37874
1.2Å	44386	33873	39109	53620	41574	43101
1.3Å	40309	36914	40913	54654	40162	43778
1.4Å	39632	32981	40205	39987	39384	42555
1.5Å	39871	37802	40188	40797	38697	41321
1.6Å	39340	39260	40112	40368	38732	42472

Table C.12: Number of simulation steps per model system / constraint for a cubic box with a side length of 19.7340 Å. Each time step corresponds to 0.5 fs. Abbreviations: $Ru^{II}OMe$ = $[Ru^{II}Py_5OMe(H_2O)]^{2+}$, $Ru^{III}OMe$ = $[Ru^{III}Py_5OMe(H_2O)]^{3+}$

d(A−H)	H_2O	$Ru^{II}OMe$	$Ru^{III}OMe$
0.9Å	33538	32542	33592
1.0Å	34498	31453	33905
1.1Å	28697	31082	28762
1.2Å	31662	31161	32810
1.3Å	29924	31933	33278
1.4Å	30428	31864	32835
1.5Å	30556	31575	32175
1.6Å	26253	30264	33664

D Zooming in on the O–O Bond Formation

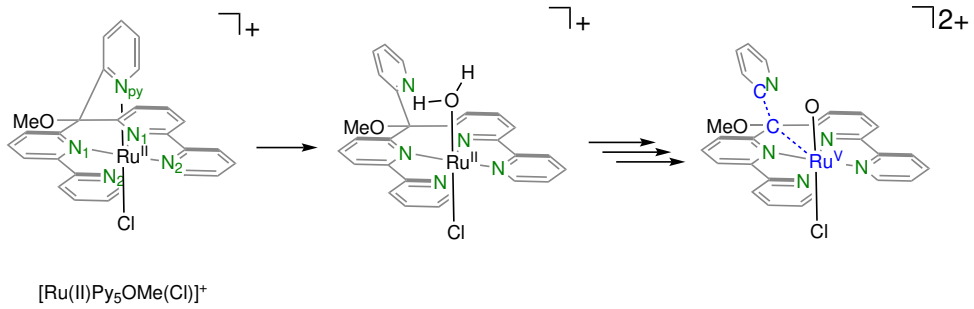


Figure D.1: Schematic representation of the catalyst as synthesized (left), after the decoordination of the pyridine (middle), and as a metal-oxo species ready for a WNA (right). The ($C_{py}C_{OMe}Ru$) angle mentioned in the discussion of the DFT optimized structures is highlighted in blue. Note the dashed line connecting Ru–C does not represent an actual bond, while the C–C of course does.

Table D.1: Comparison of gas-phase structures optimized with different general gradient approximation exchange-correlation functionals to the x-ray structure.

	{Ru(Py5OMe)}			
	d(N_{py}Ru) (Å)	d(O_wRu) (Å)	$\angle N_1 Ru N_1'^{\circ}$	$\angle N_2 Ru N_2'^{\circ}$
BLYP-D3[198,224]	2.04	2.26	94.95	109.32
BP86-D3l[224,225]	2.02	2.23	95.01	109.08
PBE-D3[192]	2.02	2.23	95.13	109.10
revPBE-D3[340]	2.02	2.27	95.23	109.49
X-ray structure[173]	2.02	2.15	95.24	108.68

D.1 Bluemoon

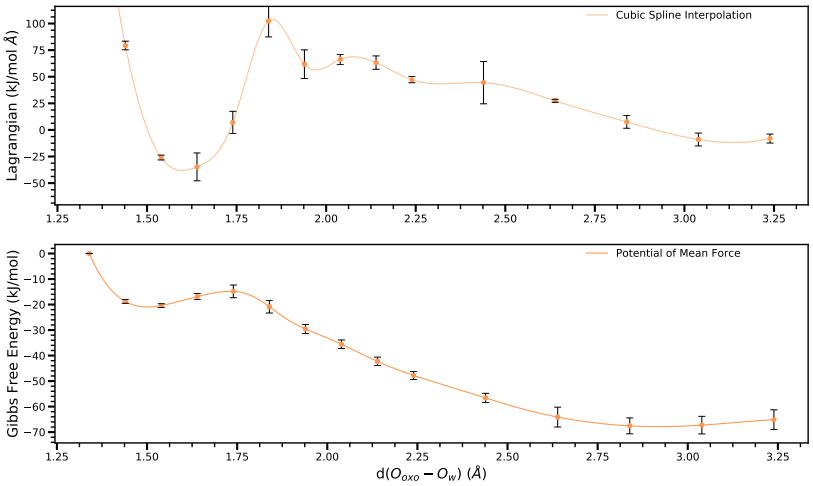


Figure D.2: Top: Exemplary force-profile of the base-assisted O–O bond formation. Bottom: Potential of mean force obtained by integration of the force-profile. The errors are obtained by block averaging over 5 blocks.

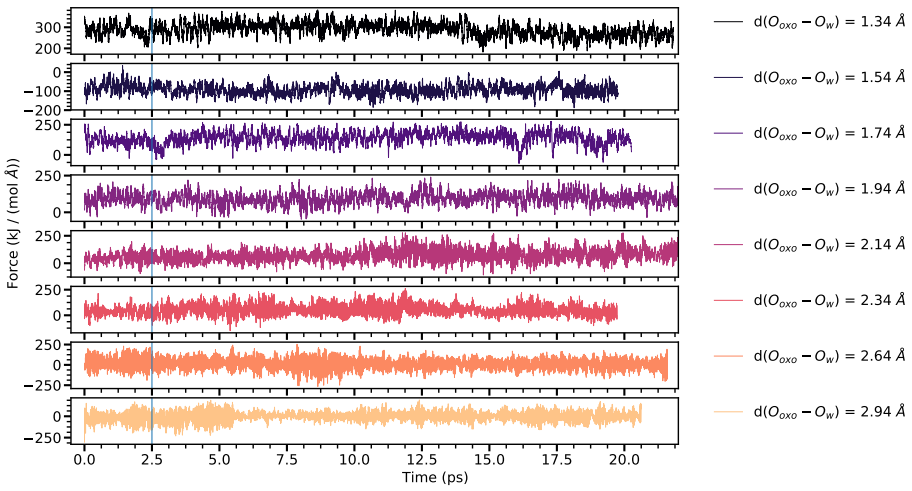


Figure D.3: Exemplary time series of the Lagrangian multiplier in case of the base-independent mechanism.

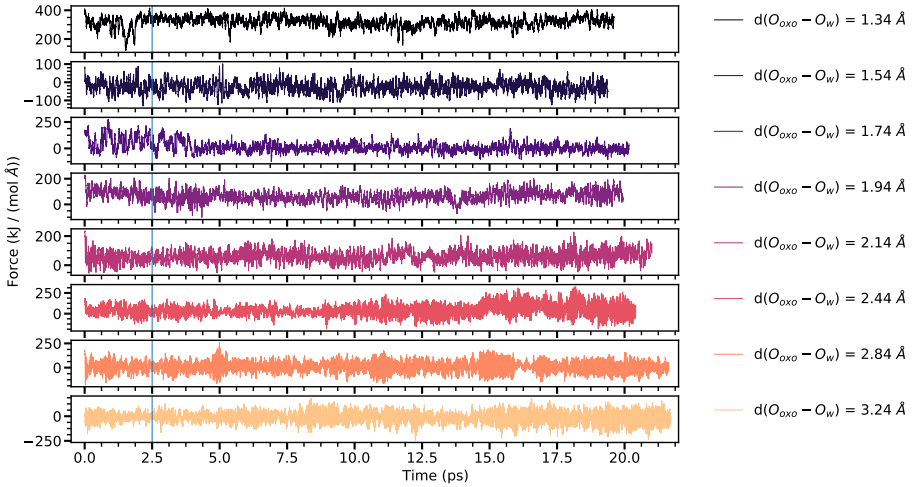


Figure D.4: Exemplary time series of the Lagrangian multiplier in case of the base-assisted mechanism.

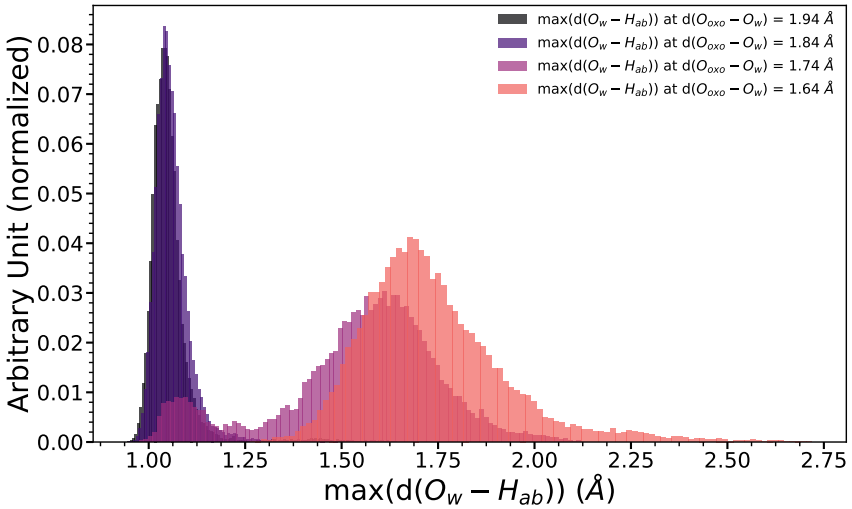


Figure D.5: Distribution of the $O_w - H_a$ distance in the region of the TS ($d(O_{oxo} - O_w)$ = 1.64 to 1.94 Å) indicated by the different colors. Note that $O_w - H_a$ distances in the rage of 1.25 Å are barely explored, the proton is either bound to the base or to the nucleophile. The area enclosed by the individual histograms is normalized to unity.

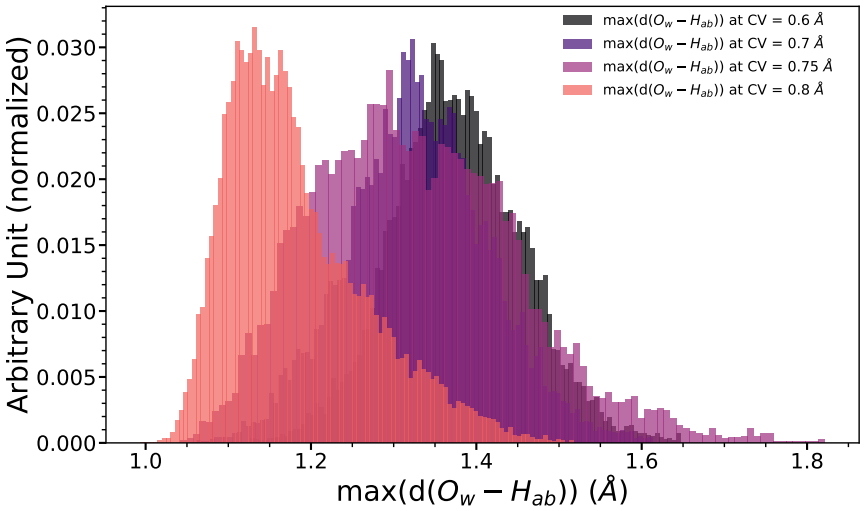


Figure D.6: Distribution of the O_w–H_a distance in the region of the proton transfer (d(O_{oxo}–O_w)-d(O_w–H_a) = 0.6 to 0.8 Å) indicated by the different colors. The area enclosed by the individual histograms is normalized to unity.

D.2 Restraining Potential

In PLUMED a restraining potential is defined as follows:

$$\sum_i \frac{1}{s_i} \kappa_i (x_i - a_i + o_i)^{e_i} \tag{9.1}$$

where κ_i is an energy constant in the internal unit of PLUMED, s_i a rescaling factor, e_i the exponent of the potential, a_i the target value and o_i the off-set.

Table D.2: Details for the applied restraining potentials in the MetaD simulation set A

	Set A type	κ	e_i	s_i	o_i
CV1	$d(O_{oxo} - O_w)$	1500	2	0.8	0.0
restraining pot.	$d(OH_a) - d(NH_b) - d(OH_b) + d(NH_a)$	500	2	0.8	0.0

Table D.3: Details for the applied restraining potentials in the MetaD simulation set B

	Set B type	κ	e_i	s_i	o_i
CV1	$d(O_{oxo} - O_w)$	1500	2	0.8	0.0
CV2	$CN(NH) - CN(OH)$	500	2	0.8	0.0

D.3 MEP Block Number

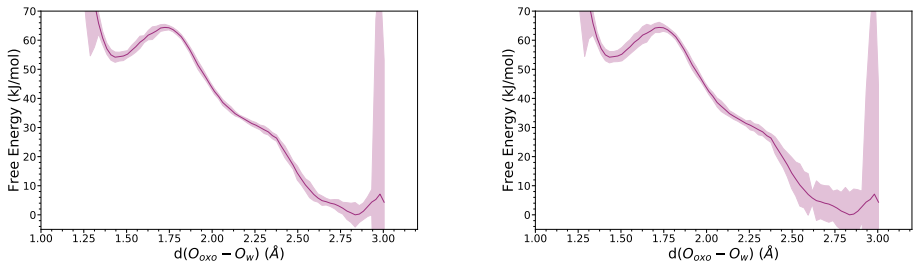


Figure D.7: Free energy profile including error range of the $O_{oxo}-O_w$ distance for {Ru(Py5OMe)} (set B). Left: number of blocks used for block averaging procedure: 10. Right: number of blocks used for block averaging procedure: 15.

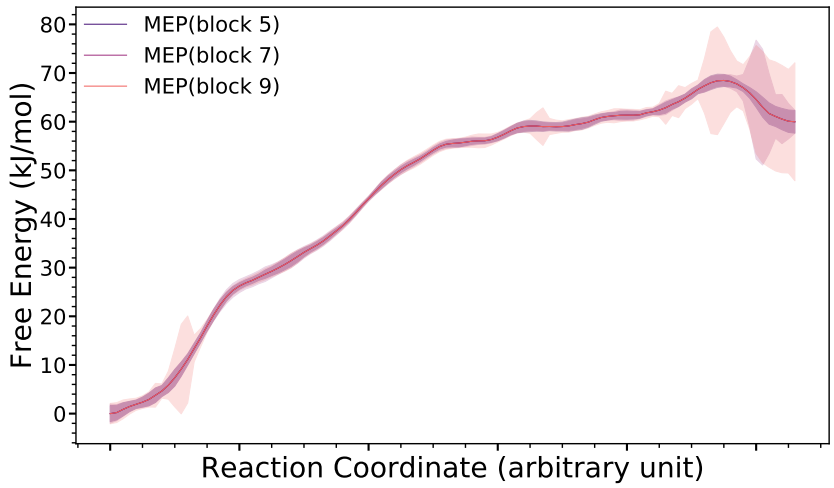


Figure D.8: MEP for the base-assisted WNA including the standard deviation obtained by different numbers of blocks.

D.4 Structural Features of the Extrema

In the Tables D.4 and D.5 key structural features are compared. In most cases, there is a good agreement among the three different methodologies. However, in particular for the MetaD simulation a shorter $O_{oxo}-O_w$ bond is found for the P state. Together with the slightly elongated min(d(O_w–H)), this suggests that the formed hydroperoxide ligand is acidic and tends to deprotonate. Another significant difference is observed for the N–O_w distance in the R state, which serves as a probe to describe the relative orientation of the nucleophile with respect to the base. In the case of the MetaD simulation this distance is increased significantly suggesting that not only states where the nucleophile is interacting with the base are sampled. The same observation is also made for min(d(N–H)) distance, which describes the

proton transfered to the base. The differences in the R state among the different methodologies are attributed to the fact that in the case of static DFT optimizations R is actually AR which as discussed in the main text is not a local minimum. The differences between *Bluemoon* and MetaD in terms of the d(N−O_w)) and d(min(N−H)) distances are caused by the fact that within the *Bluemoon* ensemble there is the possibility that at larger distances the proton not considered in the CV tends to form a hydrogen bond with the base.

Table D.4: Comparison of key structural features of the R, TS and P among the three methodologies. Since the protons of the nucleophile are indistinguishable, min(d(O−H)) is used to describe the proton remaining at the nucleophile, while max(d(O_w−H)) describes the proton being transfered to the base. Analogously, min(d(N−H)) describes the proton transfered to the base. The larger uncertainties in case of the MetaD (set B) simulations in part originate from the fact that only every 100th frame was used for the cluster analysis.

		{Ru(Py5OMe)}		
		R	TS	P
d(O_{oxo}−O_w)	static DFT	2.28	1.71	1.51
	Bluemoon	2.90 ± 0.03	1.79 ± 0.08	1.50 ± 0.05
	MetaD	2.82 ± 0.03	1.70 ± 0.06	1.46 ± 0.07
d(N−O_w)	static DFT	2.76	2.57	2.69
	Bluemoon	2.93 ± 0.18	2.63 ± 0.07	2.71 ± 0.05
	MetaD	3.46 ± 0.29	2.69 ± 0.18	2.67 ± 0.23
max(d(O_w−H))	static DFT	1.01	1.45	1.64
	Bluemoon	1.00 ± 0.03	1.59 ± 0.08	1.70 ± 0.05
	MetaD	1.02 ± 0.04	1.70 ± 0.29	2.11 ± 0.67
min(d(O_w−H))	static DFT	1.00	1.00	0.99
	Bluemoon	0.99 ± 0.02	1.00 ± 0.03	1.04 ± 0.04
	MetaD	0.99 ± 0.03	1.02 ± 0.06	1.14 ± 0.11
min(d(N−H))	static DFT	1.76	1.13	1.07
	Bluemoon	2.02 ± 0.21	1.08 ± 0.04	1.06 ± 0.03
	MetaD	2.77 ± 0.36	1.08 ± 0.06	1.18 ± 0.13
d(Ru−O_{oxo})	static DFT	1.73	1.83	1.92
	Bluemoon	1.73 ± 0.03	1.81 ± 0.04	1.92 ± 0.04
	MetaD	1.74 ± 0.03	1.84 ± 0.05	1.95 ± 0.05

Table D.5: Comparison of key structural features of the peroxo-species and deprotonated (P2). Since the protons of the nucleophile are indistinguishable min(d(O_w−H)) is used to describe the proton remaining at the nucleophile, while max(d(O_w−H)) describes the proton being transfered to the base. Analogously min(d(N−H)) describes the proton transfered to the base. All reported values belong to the simulation applying the CVs of set B.

		{Ru(Py5OMe)}	
		peroxo	deprotonated P
d(O_{oxo}−O_w)	MetaD	1.47 ± 0.07	1.47 ± 0.08
d(N−O_w)	MetaD	2.67 ± 0.13	2.72 ± 0.23
max(d(O_w−H))	MetaD	2.03 ± 0.58	2.28 ± 0.64
min(d(O_w−H))	MetaD	1.59 ± 0.11	1.04 ± 0.05
min(d(N−H))	MetaD	1.25 ± 0.29	1.75 ± 0.39
d(Ru−O_{oxo})	MetaD	1.93 ± 0.05	1.93 ± 0.05

Bibliography

[1] Böhler, M. Computational Investigation and Design Aspects of Ruthenium Py5R-type Water-Oxidation Catalysts. M.Sc. thesis, University of Zurich, 2017.

[2] Schilling, M.; Luber, S. In *Water Oxidation Catalysts*; van Eldik, R., Hubbard, C. D., Eds.; Adv. Inorg. Chem.; Academic Press, 2019; Vol. 74; pp 61–114.

[3] Huber, M.; Knutti, R. *Nat. Geosci.* **2011**, *5*, 31–36.

[4] Gudmundsson, L.; Seneviratne, S. I.; Zhang, X. *Nat. Clim. Change* **2017**, *7*, 813–816.

[5] Schiermeier, Q. *Nature* **2018**, *560*, 20–22.

[6] United Nations / Framework Convention on Climate Change(2015), - **2015**, –.

[7] Denholm, P.; O'Connell, M.; Brinkman, G.; Jorgenson, J. *National Renewable Energy Laboratory* **2015**, –.

[8] Reece, S. Y.; Hamel, J. A.; Sung, K.; Jarvi, T. D.; Esswein, A. J.; Pijpers, J. J. H.; Nocera, D. G. *Science* **2011**, *334*, 645–648.

[9] Nocera, D. G. *Acc. Chem. Res.* **2012**, *45*, 767–776.

[10] Li, L.; Duan, L.; Xu, Y.; Gorlov, M.; Hagfeldt, A.; Sun, L. *Chem. Commun.* **2010**, *46*, 7307–7309.

[11] Yang, T.; Yin, H.; Gao, L.-H.; Wang, K.-Z.; Yan, D. In *Water Oxidation Catalysts*; van Eldik, R., Hubbard, C. D., Eds.; Academic Press, 2019; Vol. 74; pp 305–341.

[12] Shaffer, D. W.; Xie, Y.; Concepcion, J. J. *Chem. Soc. Rev.* **2017**, *46*, 6170–6193.

[13] Schilling, M.; Luber, S. *Front Chem* **2018**, *6*, 100.

[14] Liu, X.; Wang, F. *Coord. Chem. Rev* **2012**, *256*, 1115–1136.

[15] Singh, A.; Spiccia, L. *Coord. Chem. Rev* **2013**, *257*, 2607–2622.

[16] Najafpour, M. M.; Renger, G.; Hołyńska, M.; Moghaddam, A. N.; Aro, E.-M.; Carpentier, R.; Nishihara, H.; Eaton-Rye, J. J.; Shen, J.-R.; Allakhverdiev, S. I. *Chem. Rev.* **2016**, *116*, 2886–2936.

[17] Tong, L.; Thummel, R. P. *Chem. Sci.* **2016**, *7*, 6591–6603.

[18] Yamamoto, M.; Tanaka, K. *ChemPlusChem* **2016**, *81*, 1028–1044.

[19] Li, J.; Güttinger, R.; Moré, R.; Song, F.; Wan, W.; Patzke, G. R. *Chem. Soc. Rev.* **2017**, *46*, 6124–6147.

[20] Dau, H.; Limberg, C.; Reier, T.; Risch, M.; Roggan, S.; Strasser, P. *ChemCatChem* **2010**, *2*, 724–761.

[21] Sala, X.; Maji, S.; Bofill, R.; García-Antón, J.; Escriche, L.; Llobet, A. *Acc. Chem. Res.* **2014**, *47*, 504–516.

[22] Blakemore, J. D.; Crabtree, R. H.; Brudvig, G. W. *Chem. Rev.* **2015**, *115*, 12974–13005.

[23] Li, Y.-Y.; Ye, K.; Siegbahn, P. E. M.; Liao, R.-Z. *ChemSusChem* **2017**, *10*, 903–911.

[24] Liao, R.-Z.; Siegbahn, P. E. M. *ChemSusChem* **2017**, *10*, 4225–4225.

[25] Vogiatzis, K. D.; Polynski, M. V.; Kirkland, J. K.; Townsend, J.; Hashemi, A.; Liu, C.; Pidko, E. A. *Chem. Rev.* **2019**, *119*, 2453–2523.

[26] Ashley, D. C.; Baik, M.-H. *ACS Catal.* **2016**, *6*, 7202–7216.

[27] Sit, P. H.-L.; Car, R.; Cohen, M. H.; Selloni, A. *Inorg. Chem.* **2011**, *50*, 10259–10267.

[28] Thom, A. J. W.; Sundstrom, E. J.; Head-Gordon, M. *Phys. Chem. Chem. Phys.* **2009**, *11*, 11297–11304.

[29] Aullón, G.; Alvarez, S. *Theor. Chem. Acc.* **2009**, *123*, 67–73.

[30] Vidossich, P.; Lledos, A. *Dalton Trans.* **2014**, *43*, 11145–11151.

[31] Reeves, Y., Kyle G.and Kanai *J. Chem. Phys.* **2014**, *141*, 024305.

[32] Hetterscheid, D. G. H.; van der Vlugt, J. I.; de Bruin, B.; Reek, J. N. H. *Angw. Chem. Int. Ed.* **2009**, *48*, 8178–8181.

[33] Nyhlén, J.; Duan, L.; Åkermark, B.; Sun, L.; Privalov, T. *Angw. Chem. Int. Ed.* **2010**, *49*, 1773–1777.

[34] Tong, L.; Duan, L.; Xu, Y.; Privalov, T.; Sun, L. *Angw. Chem. Int. Ed.* **2011**, *50*, 445–449.

[35] Scherrer, D.; Schilling, M.; Luber, S.; Fox, T.; Spingler, B.; Alberto, R.; Richmond, C. J. *Dalton Trans.* **2016**, *45*, 19361–19367.

[36] Matheu, R.; Ertem, M. Z.; Benet-Buchholz, J.; Coronado, E.; Batista, V. S.; Sala, X.; Llobet, A. *J. Am. Chem. Soc.* **2015**, *137*, 10786–10795.

[37] Govindarajan, N.; Tiwari, A.; Ensing, B.; Meijer, E. J. *Inorg. Chem.* **2018**, *57*, 13063–13066.

[38] Evangelisti, F.; Moré, R.; Hodel, F.; Luber, S.; Patzke, G. R. *J. Am. Chem. Soc.* **2015**, *137*, 11076–11084.

[39] Pushkar, Y.; Pineda-Galvan, Y.; Ravari, A. K.; Otroshchenko, T.; Hartzler, D. A. *J. Am. Chem. Soc.* **2018**, *140*, 13538–13541.

[40] Govindarajan, N.; Meijer, E. J. *Inorganics* **2019**, *7*.

[41] Zhan, S.; De Gracia Triviño, J. A.; Ahlquist, M. S. G. *Journal of the American Chemical Society* **2019**, *141*, 10247–10252.

[42] Plessow, P. *J. Chem. Theory Comput.* **2013**, *9*, 1305–1310.

[43] Henkelman, G.; Jónsson, H. *J. Chem. Phys.* **2000**, *113*, 9978–9985.

[44] Henkelman, G.; Uberuaga, B. P.; Jónsson, H. *J. Chem. Phys.* **2000**, *113*, 9901–9904.

[45] Laio, A.; Parrinello, M. *Proc Natl Acad Sci USA* **2002**, *99*, 12562–12566.

[46] Barducci, A.; Bonomi, M.; Parrinello, M. *WIREs Comput Mol Sci.* **2011**, *1*, 826–843.

[47] Barducci, A.; Bussi, G.; Parrinello, M. *Phys. Rev. Lett.* **2008**, *100*, 020603.

[48] Dama, J. F.; Rotskoff, G.; Parrinello, M.; Voth, G. A. *J. Chem. Theory Comput.* **2014**, *10*, 3626–3633.

[49] Gil-Ley, A.; Bussi, G. *J. Chem. Theory Comput.* **2015**, *11*, 1077–1085.

[50] Roos, B. *Chem. Phys.* **1980**, *48*, 157–173.

[51] Olsen, J.; Roos, B. O.; Jørgensen, P.; Jensen, H. J. A. *J. Chem. Phys.* **1988**, *89*, 2185–2192.

[52] Malmqvist, P. A.; Rendell, A.; Roos, B. O. *J. Phys. Chem.* **1990**, *94*, 5477–5482.

[53] Fleig, T.; Olsen, J.; Marian, C. M. *J. Chem. Phys.* **2001**, *114*, 4775–4790.

[54] Dongxia, M.; Li Manni, G.; Gagliardi, L. *J. Chem. Phys.* **2011**, *135*, 044128.

[55] Venturinelli Jannuzzi, S. A.; Phung, Q. M.; Domingo, A.; Formiga, A. L. B.; Pierloot, K. *Inorg. Chem.* **2016**, *55*, 5168–5179.

[56] Sparta, M.; Neese, F. *Chem. Soc. Rev.* **2014**, *43*, 5032–5041.

[57] Liakos, D. G.; Sparta, M.; Kesharwani, M. K.; Martin, J. M. L.; Neese, F. *J. Chem. Theory Comput.* **2015**, *11*, 1525–1539.

[58] Saitow, M.; Becker, U.; Riplinger, C.; F., V. E.; Neese, F. *J. Chem. Phys.* **2017**, *146*, 164105.

[59] Klamt, A. *WIREs Comput Mol Sci.* **2011**, *1*, 699–709.

[60] Cramer, C. J.; Truhlar, D. G. *Chem. Rev.* **1999**, *99*, 2161–2200.

[61] Tomasi, J. *Acc. Chem. Res.* **2004**, *112*, 184–203.

[62] Tomasi, J.; Cancés, E.; Pomelli, C. S.; Caricato, M.; Scalmani, G.; Frisch, M. J.; Cammi, R.; Basilevsky, M. V.; Chuev, G. N.; Mennucci, B. *Continuum Solvation Models in Chemical Physics*; John Wiley & Sons, Ltd, Chapter 1, pp 1–123.

[63] Skyner, R. E.; McDonagh, J. L.; Groom, C. R.; van Mourik, T.; Mitchell, J. B. O. *Phys. Chem. Chem. Phys.* **2015**, *17*, 6174–6191.

[64] Klamt, A.; Jonas, V.; Bürger, T.; Lohrenz, J. C. W. *J. Phys. Chem. A* **1998**, *102*, 5074–5085.

[65] Ho, J.; Klamt, A.; Coote, M. L. *J. Phys. Chem. A* **2010**, *114*, 13442–13444.

[66] Baik, M.-H.; Friesner, R. A. *J. Phys. Chem. A* **2002**, *106*, 7407–7412.

[67] Hodel, F. H.; Luber, S. *ACS Catal.* **2016**, *6*, 6750–6761.

[68] Hodel, F. H.; Deglmann, P.; Luber, S. *J. Chem. Theory Comput.* **2017**, *13*, 3348–3358.

[69] Vallés-Pardo, J. L.; Guijt, M. C.; Iannuzzi, M.; Joya, K. S.; de Groot, H. J. M.; Buda, F. *Chem. Phys. Chem.* **2012**, *13*, 140–146.

[70] Piccinin, S.; Sartorel, A.; Aquilanti, G.; Goldoni, A.; Bonchio, M.; Fabris, S. *Proc Natl Acad Sci USA* **2013**, *110*, 4917–4922.

[71] Zhan, S.; Mårtensson, D.; Purg, M.; Kamerlin, S. C. L.; Ahlquist, M. S. G. *Angw. Chem. Int. Ed.* **2017**, *56*, 6962–6965.

[72] Zhan, S.; Zou, R.; Ahlquist, M. S. G. *ACS Catal.* **2018**, *8*, 8642–8648.

[73] Warshel, A. *Computer Modeling of Chemical Reactions in Enzymes and Solutions*, 1st ed.; Wiley Professional; Wiley-Interscience, 1997.

[74] Wang, L.; Duan, L.; Stewart, B.; Pu, M.; Liu, J.; Privalov, T.; Sun, L. *J. Am. Chem. Soc.* **2012**, *134*, 18868–18880.

[75] Ho, J.; Coote, M. L. *Theor. Chem. Acc.* **2009**, *125*, 3.

[76] Marenich, A. V.; Ho, J.; Coote, M. L.; Cramer, C. J.; Truhlar, D. G. *Phys. Chem. Chem. Phys.* **2014**, *16*, 15068–15106.

[77] Zhan, C.-G.; Dixon, D. A. *J. Phys. Chem. A* **2001**, *105*, 11534–11540.

[78] Keith, J. A.; Grice, K. A.; Kubiak, C. P.; Carter, E. A. *J. Am. Chem. Soc.* **2013**, *135*, 15823–15829.

[79] Ho, J. *Phys. Chem. Chem. Phys.* **2015**, *17*, 2859–2868.

[80] Jaque, P.; Marenich, A. V.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem. C* **2007**, *111*, 5783–5799.

[81] Rulíšek, L. *J. Phys. Chem. C* **2013**, *117*, 16871–16877.

[82] Bím, D.; Rulíšek, L.; Srnec, M. *J. Phys. Chem. Lett.* **2016**, *7*, 7–13.

[83] Cheng, J.; Liu, X.; VandeVondele, J.; Sulpizi, M.; Sprik, M. *Acc. Chem. Res.* **2014**, *47*, 3522–3529.

[84] Blumberger, J.; Tavernelli, I.; Klein, M. L.; Sprik, M. *J. Chem. Phys.* **2006**, *124*, 064507.

[85] Sulpizi, M.; Sprik, M. *Phys. Chem. Chem. Phys.* **2008**, *10*, 5238–5249.

[86] Adriaanse, C.; Sulpizi, M.; VandeVondele, J.; Sprik, M. *J. Am. Chem. Soc.* **2009**, *131*, 6046–6047.

[87] Cheng, J.; Sulpizi, M.; Sprik, M. *J. Chem. Phys.* **2009**, *131*, 154504.

[88] Sulpizi, M.; Sprik, M. *J. Phys. Condens. Matter* **2010**, *22*, 284116.

[89] Marcus, R. A. *Rev. Mod. Phys.* **1993**, *65*, 599–610.

[90] De Meyer, T.; Ensing, B.; Rogge, S. M. J.; De Clerck, K.; Meijer, E. J.; Van Speybroeck, V. *ChemPlusChem* **2016**, *17*, 3447–3459.

[91] Hodel, F. H.; Luber, S. *J. Chem. Theory Comput.* **2017**, *13*, 974–981.

[92] Sprik, M.; Ciccotti, G. *J. Chem. Phys.* **1998**, *109*, 7737–7744.

[93] Ciccotti, G.; Ferrario, M. *Mol. Sim.* **2004**, *30*, 787–793.

[94] Ciccotti, G.; Kapral, R.; Vanden-Eijnden, E. *ChemPlusChem* **2005**, *6*, 1809–1814.

[95] Brüssel, M.; Di Dio, P. J.; Muñiz, K.; Kirchner, B. *Int. J. Mol. Sci.* **2011**, *12*, 1389–1409.

[96] Doltsinis, N. L.; Sprik, M. *Phys. Chem. Chem. Phys.* **2003**, *5*, 2612–2618.

[97] Davies, J. E.; Doltsinis, N. L.; Kirby, A. J.; Roussev, C. D.; Sprik, M. *J. Am. Chem. Soc.* **2002**, *124*, 6594–6599.

[98] Schilling, M.; Luber, S. *Inorganics* **2019**, *7*, 73.

[99] Umena, Y.; Kawakami, K.; Shen, J.-R.; Kamiya, N. *Nature* **2011**, *473*, 55–60.

[100] Siegbahn, P. E. M. *Acc. Chem. Res.* **2009**, *42*, 1871–1880.

[101] Pantazis, D. A.; Ames, W.; Cox, N.; Lubitz, W.; Neese, F. *Angw. Chem. Int. Ed.* **2012**, *51*, 9935–9940.

[102] Yano, J.; Yachandra, V. *Chem. Rev.* **2014**, *114*, 4175–4205.

[103] Vinyard, D. J.; Khan, S.; Brudvig, G. W. *Faraday Discuss.* **2015**, *185*, 37–50.

[104] Lohmiller, T.; Krewald, V.; Sedoud, A.; Rutherford, A. W.; Neese, F.; Lubitz, W.; Pantazis, D. A.; Cox, N. *J. Am. Chem. Soc.* **2017**, *139*, 14412–14424.

[105] Suga, M. et al. *Nature* **2017**, *543*, 131–135.

[106] Wang, J.; Askerka, M.; Brudvig, G. W.; Batista, V. S. *ACS Energy Lett.* **2017**, *2*, 2299–2306.

[107] Askerka, M.; Wang, J.; Brudvig, G. W.; Batista, V. S. *Biochemistry* **2014**, *53*, 6860–6862.

[108] Luber, S.; Rivalta, I.; Umena, Y.; Kawakami, K.; Shen, J.-R.; Kamiya, N.; Brudvig, G. W.; Batista, V. S. *Biochemistry* **2011**, *50*, 6308–6311.

[109] Han, Z.; Horak, K. T.; Lee, H. B.; Agapie, T. *J. Am. Chem. Soc.* **2017**, *139*, 9108–9111.

[110] Zhang, C.; Chen, C.; Dong, H.; Shen, J.-R.; Dau, H.; Zhao, J. *Science* **2015**, *348*, 690–693.

[111] Kanady, J. S.; Lin, P.-H.; Carsch, K. M.; Nielsen, R. J.; Takase, M. K.; Goddard, W. A.; Agapie, T. *J. Am. Chem. Soc.* **2014**, *136*, 14373–14376.

[112] Mukherjee, S.; Stull, J. A.; Yano, J.; Stamatatos, T. C.; Pringouri, K.; Stich, T. A.; Abboud, K. A.; Britt, R. D.; Yachandra, V. K.; Christou, G. *Proc Natl Acad Sci USA* **2012**, *109*, 2257–2262.

[113] Kanady, J. S.; Tsui, E. Y.; Day, M. W.; Agapie, T. *Science* **2011**, *333*, 733–736.

[114] McCool, N. S.; Robinson, D. M.; Sheats, J. E.; Dismukes, G. C. *J. Am. Chem. Soc.* **2011**, *133*, 11446–11449.

[115] Evangelisti, F.; Güttinger, R.; Moré, R.; Luber, S.; Patzke, G. R. *J. Am. Chem. Soc.* **2013**, *135*, 18734–18737.

[116] Song, F.; Moré, R.; Schilling, M.; Smolentsev, G.; Azzaroli, N.; Fox, T.; Luber, S.; Patzke, G. R. *J. Am. Chem. Soc.* **2017**, *139*, 14198–14208.

[117] Wang, L.-P.; Van Voorhis, T. *J. Phys. Chem. Lett.* **2011**, *2*, 2200–2204.

[118] Kwapien, K.; Piccinin, S.; Fabris, S. *J. Phys. Chem. Lett.* **2013**, *4*, 4223–4230.

[119] Li, X.; Siegbahn, P. E. M. *J. Am. Chem. Soc.* **2013**, *135*, 13804–13813.

[120] Fernando, A.; Aikens, C. M. *J. Phys. Chem. C* **2015**, *119*, 11072–11085.

[121] Beattie, J. K.; Hambley, T. W.; Klepetko, J. A.; Masters, A. F.; Turner, P. *Polyhedron* **1998**, *17*, 1343 – 1354.

[122] Smith, P. F.; Hunt, L.; Laursen, A. B.; Sagar, V.; Kaushik, S.; Calvinho, K. U. D.; Marotta, G.; Mosconi, E.; De Angelis, F.; Dismukes, G. C. *J. Am. Chem. Soc.* **2015**, *137*, 15460–15468.

[123] Nguyen, A. I.; Ziegler, M. S.; Oña-Burgos, P.; Sturzbecher-Hohne, M.; Kim, W.; Bellone, D. E.; Tilley, T. D. *J. Am. Chem. Soc.* **2015**, *137*, 12865–12872.

[124] Nguyen, A. I.; Wang, J.; Levine, D. S.; Ziegler, M. S.; Tilley, T. D. *Chem. Sci.* **2017**, *8*, 4274–4284.

[125] Zhang, B.; Sun, L. *Dalton Trans.* **2018**, *47*, 14381–14387.

[126] Brodsky, C. N.; Hadt, R. G.; Hayes, D.; Reinhart, B. J.; Li, N.; Chen, L. X.; Nocera, D. G. *Proc Natl Acad Sci USA* **2017**, *114*, 3855–3860.

[127] Nguyen, A. I.; Suess, D. L. M.; Darago, L. E.; Oyala, P. H.; Levine, D. S.; Ziegler, M. S.; Britt, R. D.; Tilley, T. D. *J. Am. Chem. Soc.* **2017**, *139*, 5579–5587.

[128] Amtawong, J.; Balcells, D.; Wilcoxen, J.; Handford, R. C.; Biggins, N.; Nguyen, A. I.; Britt, R. D.; Tilley, T. D. *J. Am. Chem. Soc.* **0**, *0*, null.

[129] Ullman, A. M.; Liu, Y.; Huynh, M.; Bediako, D. K.; Wang, H.; Anderson, B. L.; Powers, D. C.; Breen, J. J.; Abruña, H. D.; Nocera, D. G. *J. Am. Chem. Soc.* **2014**, *136*, 17681–17688.

[130] Genoni, A.; La Ganga, G.; Volpe, A.; Puntoriero, F.; Di Valentin, M.; Bonchio, M.; Natali, M.; Sartorel, A. *Faraday Discuss.* **2015**, *185*, 121–141.

[131] Hodel, F. H.; Luber, S. *ACS Catal.* **2016**, *6*, 1505–1517.

[132] Cox, N.; Retegan, M.; Neese, F.; Pantazis, D. A.; Boussac, A.; Lubitz, W. *Science* **2014**, *345*, 804–808.

[133] Schilling, M.; Hodel, F.; Luber, S. *ChemSusChem* **2017**, *10*, 4561–4569.

[134] Song, F.; Al-Ameed, K.; Schilling, M.; Fox, T.; Luber, S.; Patzke, G. R. *J. Am. Chem. Soc.* **2019**, *141*, 8846–8857.

[135] Gersten, S. W.; Samuels, G. J.; Meyer, T. J. *J. Am. Chem. Soc.* **1982**, *104*, 4029–4030.

[136] Gilbert, J. A.; Eggleston, D. S.; Murphy, W. R.; Geselowitz, D. A.; Gersten, S. W.; Hodgson, D. J.; Meyer, T. J. *J. Am. Chem. Soc.* **1985**, *107*, 3855–3864.

[137] Sala, X.; Romero, I.; Rodríguez, M.; Escriche, L.; Llobet, A. *Angw. Chem. Int. Ed.* **2009**, *48*, 2842–2852.

[138] Romain, S.; Vigara, L.; Llobet, A. *Acc. Chem. Res.* **2009**, *42*, 1944–1953.

[139] Cao, R.; Lai, W.; Du, P. *Energy Environ. Sci.* **2012**, *5*, 8134–8157.

[140] Kärkäs, M. D.; Verho, O.; Johnston, E. V.; Åkermark, B. *Chem. Rev.* **2014**, *114*, 11863–12001.

[141] Zeng, Q.; Lewis, F. W.; Harwood, L. M.; Hartl, F. *Coord. Chem. Rev* **2015**, *304-305*, 88–101.

[142] Kärkäs, M. D.; Åkermark, B. *Chem. Rev.* **2016**, *16*, 940–963.

[143] Kamdar, J. M.; Grotjahn, D. B. *Molecules* **2019**, *24*.

[144] Duan, L.; Bozoglian, F.; Mandal, S.; Stewart, B.; Privalov, T.; Llobet, A.; Sun, L. *Nat. Chem.* **2012**, *4*, 418–423.

[145] Duan, L.; Xu, Y.; Zhang, P.; Wang, M.; Sun, L. *Inorg. Chem.* **2010**, *49*, 209–215.

[146] Duan, L.; Araujo, C. M.; Ahlquist, M. S.; Sun, L. *Proc Natl Acad Sci USA* **2012**,

[147] Duan, L.; Wang, L.; Inge, A. K.; Fischer, A.; Zou, X.; Sun, L. *Inorg. Chem.* **2013**, *52*, 7844–7852.

[148] Wang, L.; Duan, L.; Wang, Y.; Ahlquist, M. S. G.; Sun, L. *Chem. Commun.* **2014**, *50*, 12947–12950.

[149] Staehle, R.; Tong, L.; Wang, L.; Duan, L.; Fischer, A.; Ahlquist, M. S. G.; Sun, L.; Rau, S. *Inorg. Chem.* **2014**, *53*, 1307–1319.

[150] Fan, T.; Duan, L.; Huang, P.; Chen, H.; Daniel, Q.; Ahlquist, M. S. G.; Sun, L. *ACS Catal.* **2017**, *7*, 2956–2966.

[151] Xie, Y.; Shaffer, D. W.; Concepcion, J. J. *Inorg. Chem.* **2018**, *57*, 10533–10542.

[152] Zhang, G.; Chen, K.; Chen, H.; Yao, J.; Shaik, S. *Inorg. Chem.* **2013**, *52*, 5088–5096.

[153] Jarvis, E. A.; Lee, B.; Neddenriep, B.; Shoemaker, W. *Chem. Phys.* **2013**, *417*, 8–16.

[154] Kang, R.; Chen, K.; Yao, J.; Shaik, S.; Chen, H. *Inorg. Chem.* **2014**, *53*, 7130–7136.

[155] Asaduzzaman, A. M.; Wasylenko, D.; Berlinguette, C. P.; Schreckenbach, G. *J. Phys. Chem. C* **2015**, *119*, 242–250.

[156] Okamura, M.; Masaoka, S. *Chem. Asian J.* **2015**, *10*, 306–315.

[157] Hessels, J.; Detz, R. J.; Koper, M. T. M.; Reek, J. N. H. *Chem. Eur. J.* **2017**, *23*, 16413–16418.

[158] Richmond, C. J.; Escayola, S.; Poater, A. *Eur. J. Inorg. Chem.* **2019**, *2019*, 2101–2108.

[159] Duan, L.; Fischer, A.; Xu, Y.; Sun, L. *J. Am. Chem. Soc.* **2009**, *131*, 10397–10399.

[160] Matheu, R.; Ertem, M. Z.; Gimbert-Suriñach, C.; Benet-Buchholz, J.; Sala, X.; Llobet, A. *ACS Catal.* **2017**, *7*, 6525–6532.

[161] Matheu, R.; Ertem, M. Z.; Pipelier, M.; Lebreton, J.; Dubreuil, D.; Benet-Buchholz, J.; Sala, X.; Tessier, A.; Llobet, A. *ACS Catal.* **2018**, *8*, 2039–2048.

[162] Wang, Y.; Zhan, S.; Ahlquist, M. S. G. *Organometallics* **2019**, *38*, 1264–1268.

[163] Zong, R.; Thummel, R. P. *J. Am. Chem. Soc.* **2005**, *127*, 12802–12803.

[164] Tseng, H.-W.; Zong, R.; Muckerman, J. T.; Thummel, R. *Inorg. Chem.* **2008**, *47*, 11763–11773.

[165] Polyansky, D. E.; Muckerman, J. T.; Rochford, J.; Zong, R.; Thummel, R. P.; Fujita, E. *J. Am. Chem. Soc.* **2011**, *133*, 14649–14665.

[166] Badiei, Y. M.; Polyansky, D. E.; Muckerman, J. T.; Szalda, D. J.; Haberdar, R.; Zong, R.; Thummel, R. P.; Fujita, E. *Inorg. Chem.* **2013**, *52*, 8845–8850.

[167] Muckerman, J. T.; Kowalczyk, M.; Badiei, Y. M.; Polyansky, D. E.; Concepcion, J. J.; Zong, R.; Thummel, R. P.; Fujita, E. *Inorg. Chem.* **2014**, *53*, 6904–6913.

[168] Tong, L.; Zong, R.; Zhou, R.; Kaveevivitchai, N.; Zhang, G.; Thummel, R. P. *Faraday Discuss.* **2015**, *185*, 87–104.

[169] Moonshiram, D.; Pineda-Galvan, Y.; Erdman, D.; Palenik, M.; Zong, R.; Thummel, R.; Pushkar, Y. *J. Am. Chem. Soc.* **2016**, *138*, 15605–15616.

[170] Radaram, B.; Ivie, J. A.; Singh, W. M.; Grudzien, R. M.; Reibenspies, J. H.; Webster, C. E.; Zhao, X. *Inorg. Chem.* **2011**, *50*, 10564–10571.

[171] Vennampalli, M.; Liang, G.; Webster, C. E.; Zhao, X. *Eur. J. Inorg. Chem.* **2014**, *2014*, 715–721.

[172] Gil-Sepulcre, M.; Axelson, J. C.; Aguiló, J.; Solà-Hernández, L.; Francàs, L.; Poater, A.; Blancafort, L.; Benet-Buchholz, J.; Guirado, G.; Escriche, L.; Llobet, A.; Bofill, R.; Sala, X. *Inorg. Chem.* **2016**, *55*, 11216–11229.

[173] Gil-Sepulcre, M.; Böhler, M.; Schilling, M.; Bozoglian, F.; Bachmann, C.; Scherrer, D.; Fox, T.; Spingler, B.; Gimbert-Suriñach, C.; Alberto, R.; Bofill, R.; Sala, X.; Luber, S.; Richmond, C. J.; Llobet, A. *ChemSusChem* **2017**, *10*, 4517–4525.

[174] Schilling, M.; Böhler, M.; Luber, S. *Dalton Trans.* **2018**, *47*, 10480–10490.

[175] Schilling, M.; Patzke, G. R.; Hutter, J.; Luber, S. *J. Phys. Chem. C* **2016**, *120*, 7966–7975.

[176] Durand, D. J.; Fey, N. *Chem. Rev.* **2019**, *119*, 6561–6594.

[177] Sprik, M. *Faraday Discuss.* **1998**, *110*, 437–445.

[178] Kiliç, M.; Ensing, B. *Phys. Chem. Chem. Phys.* **2014**, *16*, 18993–19000.

[179] Bankura, A.; Klein, M. L.; Carnevale, V. *Chem. Phys.* **2013**, *422*, 156–164.

[180] Sinha, V.; Govindarajan, N.; de Bruin, B.; Meijer, E. J. *ACS Catal.* **2018**, *8*, 6908–6913.

[181] Chandler, D. *Introduction to Modern Statistical Mechanics*; Oxford University Press, Inc., 1987.

[182] Bernasconi, L.; Baerends, E. J.; Sprik, M. *J. Phys. Chem. B* **2006**, *110*, 11444–11453.

[183] Liu, X.; Lu, X.; Wang, R.; Zhou, H. *J. Phys. Chem. A* **2010**, *114*, 12914–12917.

[184] Chen, Y.-L.; Doltsinis, N. L.; Hider, R. C.; Barlow, D. J. *J. Phys. Chem. Lett.* **2012**, *3*, 2980–2985.

[185] Ivanov, I.; Chen, B.; Raugei, S.; Klein, M. L. *J. Phys. Chem. B* **2006**, *110*, 6365–6371.

[186] Marcos-Alcalde, I. n.; Setoain, J.; Mendieta-Moreno, J. I.; Mendieta, J.; Gómez-Puertas, P. *Bioinformatics* **2015**, *31*, 3853–3855.

[187] Born, M.; Oppenheimer, R. *Ann. Phys.* **1927**, *389*, 457–484.

[188] Hohenberg, P.; Kohn, W. *Phys. Rev.* **1964**, *136*, B864–B871.

[189] Kohn, W.; Sham, L. J. *Phys. Rev.* **1965**, *140*, A1133–A1138.

[190] Perdew, J. P.; Wang, Y. *Phys. Rev. B* **1992**, *45*, 13244–13249.

[191] Vosko, S. H.; Wilk, L.; Nusair, M. *Can. J. Phys.* **1980**, *58*, 1200–1211.

[192] Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1996**, *77*, 3865–3868.

[193] Perdew, J. P.; Burke, K.; Wang, Y. *Phys. Rev. B* **1996**, *54*, 16533–16539.

[194] Henderson, T. M.; Janesko, B. G.; Scuseria, G. E. *J. Phys. Chem. A* **2008**, *112*, 12530–12542.

[195] Goerigk, L.; Mehta, N. *Phys. Chem. Chem. Phys.* **2019**, *72*, 563–573.

[196] Goerigk, L.; Hansen, A.; Bauer, C.; Ehrlich, S.; Najibi, A.; Grimme, S. *Phys. Chem. Chem. Phys.* **2017**, *19*, 32184–32215.

[197] Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648–5652.

[198] Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785–789.

[199] Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. *J. Chem. Phys.* **2010**, *132*, 154104.

[200] Grimme, S. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2011**, *1*, 211–228.

[201] Grimme, S.; Ehrlich, S.; Goerigk, L. *J. Comput. Chem.* **2011**, *32*, 1456–1465.

[202] Vydrov, O. A.; Van Voorhis, T. *J. Chem. Phys.* **2010**, *133*, 244103.

[203] Bučko, T. c. v.; Lebègue, S.; Hafner, J.; Ángyán, J. G. *Phys. Rev. B* **2013**, *87*, 064110.

[204] Klamt, A.; Schüürmann, G. *J. Chem. Soc., Perkin Trans. 2* **1993**, 799–805.

[205] Schäfer, A.; Klamt, A.; Sattel, D.; Lohrenz, J. C. W.; Eckert, F. *Phys. Chem. Chem. Phys.* **2000**, *2*, 2187–2193.

[206] Ahlrichs, R.; Bär, M.; Häser, M.; Horn, H.; Kölmel, C. *Chem. Phys. Lett.* **1989**, *162*, 165–169.

[207] Verlet, L. *Phys. Rev.* **1967**, *159*, 98–103.

[208] Swope, W. C.; Andersen, H. C.; Berens, P. H.; Wilson, K. R. *J. Chem. Phys.* **1982**, *76*, 637–649.

[209] Gunsteren, W. F. V.; Berendsen, H. J. C. *Mol. Simul.* **1988**, *1*, 173–185.

[210] Nosé, S. *J. Chem. Phys.* **1984**, *81*, 511–519.

[211] Nosé, S. *Mol. Phys.* **1984**, *52*, 255–268.

[212] Car, R.; Parrinello, M. *Phys. Rev. Lett.* **1985**, *55*, 2471–2474.

[213] Elber, R. *J. Chem. Phys.* **2016**, *144*, 060901.

[214] Torrie, G.; Valleau, J. *J. Comput. Phys.* **1977**, *23*, 187–199.

[215] Yang, Y. I.; Shao, Q.; Zhang, J.; Yang, L.; Gao, Y. Q. *J. Chem. Phys.* **2019**, *151*, 070902.

[216] Palazzesi, F.; Valsson, O.; Parrinello, M. *J. Phys. Chem. Lett.* **2017**, *8*, 4752–4756.

[217] Sultan, M. M.; Pande, V. S. *J. Chem. Phys.* **2018**, *149*, 094106.

[218] den Otter, W. K. *J. Chem. Theory Comput.* **2013**, *9*, 3861–3865.

[219] Komeiji, Y. *Chem-Bio Inf. J.* **2007**, *7*, 12–23.

[220] Laio, A.; Gervasio, F. L. *Rep. Prog. Phys.* **2008**, *71*, 126601.

[221] Sutto, L.; Marsili, S.; Gervasio, F. L. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2012**, *2*, 771–779.

[222] Valsson, O.; Tiwary, P.; Parrinello, M. *Annu. Rev. Phys. Chem.* **2016**, *67*, 159–184.

[223] Díaz Leines, G.; Ensing, B. *Phys. Rev. Lett.* **2012**, *109*, 020601.

[224] Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098–3100.

[225] Perdew, J. P. *Phys. Rev. B* **1986**, *33*, 8822–8824.

[226] Weigend, F.; Ahlrichs, R. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297–3305.

[227] Deglmann, P.; Furche, F. *J. Chem. Phys.* **2002**, *117*, 9535–9538.

[228] Deglmann, P.; Furche, F.; Ahlrichs, R. *Chem. Phys. Lett.* **2002**, *362*, 511–518.

[229] Eichkorn, K.; Treutler, O.; Öhm, H.; Häser, M.; Ahlrichs, R. *Chem. Phys. Lett.* **1995**, *240*, 283–290.

[230] Eichkorn, K.; Weigend, F.; Treutler, O.; Ahlrichs, R. *Theor. Chem. Acc.* **1997**, *97*, 119–124.

[231] Weigend, F. *Phys. Chem. Chem. Phys.* **2006**, *8*, 1057–1065.

[232] Sierka, M.; Hogekamp, A.; Ahlrichs, R. *J. Chem. Phys.* **2003**, *118*, 9136–9148.

[233] Bill, E.; Bothe, E.; Chaudhuri, P.; Chlopek, K.; Herebian, D.; Kokatam, S.; Ray, K.; Weyhermüller, T.; Neese, F.; Wieghardt, K. *Chem. Eur. J.* **2005**, *11*, 204–224.

[234] Ames, W.; Pantazis, D. A.; Krewald, V.; Cox, N.; Messinger, J.; Lubitz, W.; Neese, F. *J. Am. Chem. Soc.* **2011**, *133*, 19743–19757.

[235] Orio, M.; Pantazis, D. A.; Neese, F. *Photosynth Res* **2009**, *102*, 443–453.

[236] Neese, F. *JBIC, J. Biol. Inorg. Chem.* **2006**, *11*, 702–711.

[237] Neese, F. *Coord. Chem. Rev* **2009**, *253*, 526–563, Theory and Computing in Contemporary Coordination Chemistry.

[238] Neese, F.; Ames, W.; Christian, G.; Kampa, M.; Liakos, D. G.; Pantazis, D. A.; Roemelt, M.; Surawatanawong, P.; Shengfa, Y. In *Theoretical and Computational Inorganic Chemistry*; van Eldik, R., Harvey, J., Eds.; Adv. Inorg. Chem.; Academic Press, 2010; Vol. 62; pp 301–349.

[239] Kang, R.; Yao, J.; Chen, H. *J. Chem. Theory Comput.* **2013**, *9*, 1872–1879.

[240] VandeVondele, J.; Krack, M.; Mohamed, F.; Parrinello, M.; Chassaing, T.; Hutter, J. *Comput. Phys. Commun.* **2005**, *167*, 103–128.

[241] CP2K Program Package.

[242] Goedecker, S.; Teter, M.; Hutter, J. *Phys. Rev. B* **1996**, *54*, 1703–1710.

[243] Hartwigsen, C.; Goedecker, S.; Hutter, J. *Phys. Rev. B* **1998**, *58*, 3641–3662.

[244] Guidon, M.; Hutter, J.; VandeVondele, J. *J. Chem. Theory Comput.* **2010**, *6*, 2348–2364.

[245] Thomsen, J. M.; Huang, D. L.; Crabtree, R. H.; Brudvig, G. W. *Dalton Trans.* **2015**, *44*, 12452–12472.

[246] Kondo, M.; Masaoka, S. *Chem. Lett.* **2016**, *45*, 1220–1231.

[247] Dogutan, D. K.; Stoian, S. A.; McGuire, R.; Schwalbe, M.; Teets, T. S.; Nocera, D. G. *J. Am. Chem. Soc.* **2011**, *133*, 131–140.

[248] Rosenthal, J.; Nocera, D. G. *Acc. Chem. Res.* **2007**, *40*, 543–553.

[249] Lai, W.; Cao, R.; Dong, G.; Shaik, S.; Yao, J.; Chen, H. *J. Phys. Chem. Lett.* **2012**, *3*, 2315–2319.

[250] Lei, H.; Han, A.; Li, F.; Zhang, M.; Han, Y.; Du, P.; Lai, W.; Cao, R. *Phys. Chem. Chem. Phys.* **2014**, *16*, 1883–1893.

[251] Ertem, M. Z.; Cramer, C. J. *Dalton Trans.* **2012**, *41*, 12213–12219.

[252] Ganguly, S.; Renz, D.; Giles, L. J.; Gagnon, K. J.; McCormick, L. J.; Conradie, J.; Sarangi, R.; Ghosh, A. *Inorg. Chem.* **2017**, *56*, 14788–14800.

[253] Wasylenko, D. J.; Ganesamoorthy, C.; Borau-Garcia, J.; Berlinguette, C. P. *Chem. Commun.* **2011**, *47*, 4249–4251.

[254] Crandell, D. W.; Ghosh, S.; Berlinguette, C. P.; Baik, M.-H. *ChemSusChem* **2015**, *8*, 844–852.

[255] Younus, H. A.; Ahmad, N.; Chughtai, A. H.; Vandichel, M.; Busch, M.; Van Hecke, K.; Yusubov, M.; Song, S.; Verpoort, F. *ChemSusChem* **2017**, *10*, 862–875.

[256] Ishizuka, T.; Watanabe, A.; Kotani, H.; Hong, D.; Satonaka, K.; Wada, T.; Shiota, Y.; Yoshizawa, K.; Ohara, K.; Yamaguchi, K.; Kato, S.; Fukuzumi, S.; Kojima, T. *Inorg. Chem.* **2016**, *55*, 1154–1164.

[257] Fukuzumi, S.; Mandal, S.; Mase, K.; Ohkubo, K.; Park, H.; Benet-Buchholz, J.; Nam, W.; Llobet, A. *J. Am. Chem. Soc.* **2012**, *134*, 9906–9909.

[258] Rigsby, M. L.; Mandal, S.; Nam, W.; Spencer, L. C.; Llobet, A.; Stahl, S. S. *Chem. Sci.* **2012**, *3*, 3058–3062.

[259] Gimbert-Suriãch, C.; Moonshiram, D.; Francàs, L.; Planas, N.; Bernales, V.; Bozoglian, F.; Guda, A.; Mognon, L.; López, I.; Hoque, M. A.; Gagliardi, L.; Cramer, C. J.; Llobet, A. *J. Am. Chem. Soc.* **2016**, *138*, 15291–15294.

[260] Mandal, S.; Shikano, S.; Yamada, Y.; Lee, Y.-M.; Nam, W.; Llobet, A.; Fukuzumi, S. *J. Am. Chem. Soc.* **2013**, *135*, 15294–15297.

[261] Di Giovanni, C.; Gimbert-Suriñach, C.; Nippe, M.; Benet-Buchholz, J.; Long, J. R.; Sala, X.; Llobet, A. *Chem. Eur. J.* **2016**, *22*, 361–369.

[262] Kanan, M. W.; Nocera, D. G. *Science* **2008**, *321*, 1072–1075.

[263] Smith, P. F.; Kaplan, C.; Sheats, J. E.; Robinson, D. M.; McCool, N. S.; Mezle, N.; Dismukes, G. C. *Inorg. Chem.* **2014**, *53*, 2113–2121.

[264] Mattioli, G.; Giannozzi, P.; Amore Bonapasta, A.; Guidoni, L. *J. Am. Chem. Soc.* **2013**, *135*, 15353–15363.

[265] Soriano-López, J.; Musaev, D.; L. Hill, C.; Galan Mascaros, J.; J. Carbó, J.; Poblet, J. *J. Catal.* **2017**, *350*, 56–63.

[266] Philips, J. J.; Hudspeth, M. A.; Browne, P. M.; Peralta, J. E. *Chem. Phys. Lett.* **2010**, *495*, 146–150.

[267] Narzi, D.; Bovi, D.; Guidoni, L. *Proc Natl Acad Sci USA* **2014**, *111*, 8723–8728.

[268] Nørskov, J. K.; Rossmeisl, J.; Logadottir, A.; Lindqvist, L.; Kitchin, J. R.; Bligaard, T.; Jónsson, H. *J. Phys. Chem. B* **2004**, *108*, 17886–17892.

[269] Piccinin, S.; Fabris, S. *Phys. Chem. Chem. Phys.* **2011**, *13*, 7666–7674.

[270] Valdés, A.; Kroes, G.-J. *J. Phys. Chem. C* **2010**, *114*, 1701–1708.

[271] Stull, D. R.; Prophet, H. *JANAF Thermochemical Tables*; United States Department of Commerce, National Bureau of Standards, 1971.

[272] Theilacker, K.; Buhrke, D.; Kaupp, M. *J. Chem. Theory Comput.* **2015**, *11*, 111–121.

[273] Ehrhardt, C.; Ahlrichs, R. *Theor. Chem. Acc.* **1985**, *68*, 231–245.

[274] Halgren, T. A.; Lipscomb, W. N. *Chem. Phys. Lett.* **1977**, *49*, 225–232.

[275] Elber, R.; Karplus, M. *Chem. Phys. Lett.* **1987**, *139*, 375–380.

[276] Mills, G.; Jónsson, H. *Phys. Rev. Lett.* **1994**, *72*, 1124–1127.

[277] W. Humphrey, K. S., A. Dalke *J Mol Graph* **1996**, *14*, 33–38.

[278] Legault, C. Y. CYLView.

[279] Isobe, H.; Shoji, M.; Shen, J.-R.; Yamaguchi, K. *Inorg. Chem.* **2015**, *55*, 502–511.

[280] Krewald, V.; Retegan, M.; Neese, F.; Lubitz, W.; Pantazis, D. A.; Cox, N. *Inorg. Chem.* **2015**, *55*, 488–501.

[281] Capone, M.; Bovi, D.; Narzi, D.; Guidoni, L. *Biochemistry* **2015**, *54*, 6439–6442.

[282] Retegan, M.; Krewald, V.; Mamedov, F.; Neese, F.; Lubitz, W.; Cox, N.; Pantazis, D. A. *Chem. Sci.* **2016**, *7*, 72–84.

[283] Amin, M.; Vogt, L.; Vassiliev, S.; Rivalta, I.; Sultan, M. M.; Bruce, D.; Brudvig, G. W.; Batista, V. S.; Gunner, M. *J. Phys. Chem. B* **2013**, *117*, 6217–6226.

[284] Guo, Y.; Li, H.; He, L.-L.; Zhao, D.-X.; Gong, L.-D.; Yang, Z.-Z. *Phys. Chem. Chem. Phys.* **2017**, *19*, 13909–13923.

[285] Siegbahn, P. E. *Phys. Chem. Chem. Phys.* **2012**, *14*, 4849–4856.

[286] Soriano-López, J.; Musaev, D. G.; Hill, C. L.; Galán-Mascarós, J. R.; Carbó, J. J.; Poblet, J. M. *J. Catal.* **2017**, *350*, 56–63.

[287] Crandell, D. W.; Ghosh, S.; Berlinguette, C. P.; Baik, M.-H. *ChemSusChem* **2015**, *8*, 844–852.

[288] Winkler, J. R.; Gray, H. B. In *Molecular Electronic Structures of Transition Metal Complexes I*; Mingos, D. M. P., Day, P., Dahl, J. P., Eds.; Structure and Bonding; Springer-Verlag Berlin Heidelberg, 2012; Vol. 142; pp 17–28.

[289] Reiher, M.; Sellmann, D.; Hess, B. A. *Theor. Chem. Acc.* **2001**, *106*, 379–392.

[290] Reiher, M.; Kirchner, B. *J. Phys. Chem. A* **2003**, *107*, 4141–4146.

[291] Klamt, A.; Diedenhofen, M. *J. Phys. Chem. A* **2015**, *119*, 5439–5445.

[292] Zhan, S.; Mårtensson, D.; Purg, M.; Kamerlin, S. C. L.; Ahlquist, M. S. G. *Angw. Chem. Int. Ed.* **2017**, *56*, 6962–6965.

[293] Vilella, L.; Vidossich, P.; Balcells, D.; Lledos, A. *Dalton Trans.* **2011**, *40*, 11241–11247.

[294] Tissandier, M. D.; Cowen, K. A.; Feng, W. Y.; Gundlach, E.; Cohen, M. H.; Earhart, A. D.; Coe, J. V.; Tuttle, T. R. *J. Phys. Chem. A* **1998**, *102*, 7787–7794.

[295] Kelly, C. P.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem. B* **2006**, *110*, 16066–16081.

[296] Daniel, Q.; Huang, P.; Fan, T.; Wang, Y.; Duan, L.; Wang, L.; Li, F.; Rinkevicius, Z.; Mamedov, F.; Ahlquist, M. S.; Styring, S.; Sun, L. *Coord. Chem. Rev* **2017**, *346*, 206–215.

[297] Yan, L.; Zong, R.; Pushkar, Y. *J. Catal.* **2015**, *330*, 255–260.

[298] De Vito, D.; Sidorenkova, H.; Rotzinger, F. P.; Weber, J.; Merbach, A. E. *Inorg. Chem.* **2000**, *39*, 5547–5552.

[299] Sala, X.; Ertem, M. Z.; Vigara, L.; Todorova, T. K.; Chen, W.; Rocha, R. C.; Aquilante, F.; Cramer, C. J.; Gagliardi, L.; Llobet, A. *Angw. Chem. Int. Ed.* **2010**, *49*, 7745–7747.

[300] Sabatier, P. *La Catalyse En Chimie Organique*; Librairie Polytechnique: Paris et Liège, 1920.

[301] Öğretir, C.; Özöğüt, D.; Yarligan, S.; Arslan, T. **2006**, *759*, 73–78.

[302] Taagepera, M.; Henderson, W. G.; Brownlee, R. T. C.; Beauchamp, J. L.; Holtz, D.; Taft, R. W. *J. Am. Chem. Soc.* **1972**, *94*, 1369–1370.

[303] Abdel-Magied, A. F.; Shatskiy, A.; Liao, R. Z.; Laine, T. M.; Arafa, W. A.; Siegbahn, P. E.; Kärkäs, M. D.; Åkermark, B.; Johnston, E. V. *ChemSusChem* **2016**, *9*, 3448–3456.

[304] Shields, G. S.; Seybold, P. G. *CRC Handbook of Chemistry and Physics*; CRC Press, Taylor & Francis Group, NW, 2014.

[305] Bayly, C. I.; Cieplak, P.; Cornell, W.; Kollman, P. A. *J. Phys. Chem.* **1993**, *97*, 10269–10280.

[306] am Busch, M. S.; Knapp, E.-W. *ChemPlusChem* **2004**, *5*, 1513–1522.

[307] Galstyan, G.; Knapp, E.-W. *J. Comput. Chem.* **2015**, *36*, 69–78.

[308] Simonson, T.; Carlsson, J.; Case, D. A. *J. Am. Chem. Soc.* **2004**, *126*, 4167–4180.

[309] Meyer, T.; Kieseritzky, G.; Knapp, E.-W. *Proteins* **2011**, *79*, 3320–3332.

[310] Klamt, A.; Eckert, F.; Diedenhofen, M.; Beck, M. E. *J. Phys. Chem. A* **2003**, *107*, 9380–9386.

[311] Nielsen, J. E.; Gunner, M. R.; García-Moreno E., B. *Proteins* **2011**, *79*, 3249–3259.

[312] Alexov, E.; Mehler, E. L.; Baker, N.; M. Baptista, A.; Huang, Y.; Milletti, F.; Erik Nielsen, J.; Farrell, D.; Carstensen, T.; Olsson, M. H. M.; Shen, J. K.; Warwicker, J.; Williams, S.; Word, J. M. *Proteins* **2011**, *79*, 3260–3275.

[313] Gunner, M.; Baker, N. In *Computational Approaches for Studying Enzyme Mechanism Part B*; Voth, G. A., Ed.; Methods in Enzymology; Academic Press, 2016; Vol. 578; pp 1–20.

[314] Sprik, M. *Chem. Phys.* **2000**, *258*, 139–150.

[315] Mahoney, M. W.; Jorgensen, W. L. *J. Chem. Phys.* **2000**, *112*, 8910–8922.

[316] VandeVondele, J.; Mohamed, F.; Krack, M.; Hutter, J.; Sprik, M.; Parrinello, M. *J. Chem. Phys.* **2005**, *122*, 014515.

[317] Braude, E. A.; Nachod, F. C. *Determination of Organic Structures by Physical Methods*; Academic Press., Inc., 1955; Vol. 1.

[318] Flyvbjerg, H.; Petersen, H. G. *The Journal of Chemical Physics* **1989**, *91*, 461–466.

[319] R. A. Robinson, R. G. B., M. Paabo *J Res Natl Bur Stand* **1969**, *73A*, 299–308.

[320] Delgado, R.; Silva, J. D.; Amorim, M.; Cabral, M.; Chaves, S.; Costa, J. *Anal. Chim. Acta* **1991**, *245*, 271–282.

[321] Krężel, A.; Bal, W. *J. Inorg. Biochem.* **2004**, *98*, 161–166.

[322] Lide, D. R. *CRC Handbook of Chemistry and Physics*; CRC Press, Boca Raton, FL, 2005.

[323] Sala, O.; Lüthi, H. P.; Togni, A.; Iannuzzi, M.; Hutter, J. *J. Comput. Chem.* **2015**, *36*, 785–794.

[324] Sala, O.; Santschi, N.; Jungen, S.; Lüthi, H. P.; Iannuzzi, M.; Hauser, N.; Togni, A. *Chem. Eur. J.* **2016**, *22*, 1704–1713.

[325] Liu, L.; Lukose, B.; Ensing, B. *ACS Catal.* **2018**, *8*, 3376–3381.

[326] Shao, Y.; de Ruiter, J. M.; de Groot, H. J. M.; Buda, F. *J. Phys. Chem. C* **2019**, *123*, 21403–21414.

[327] VandeVondele, J.; Hutter, J. *J. Chem. Phys.* **2007**, *127*, 114105.

[328] Koitz, R.; Iannuzzi, M.; Hutter, J. *J. Phys. Chem. C* **2015**, *119*, 4023–4030.

[329] Koitz, R.; Hutter, J.; Iannuzzi, M. *2d Mater.* **2016**, *3*, 025026.

[330] Spreafico, C.; Schiffmann, F.; VandeVondele, J. *J. Phys. Chem. C* **2014**, *118*, 6251–6260.

[331] Tribello, G. A.; Bonomi, M.; Branduardi, D.; Camilloni, C.; Bussi, G. *Comput. Phys. Commun.* **2014**, *185*, 604–613.

[332] Raiteri, P.; Laio, A.; Gervasio, F. L.; Micheletti, C.; Parrinello, M. *J. Phys. Chem. B* **2006**, *110*, 3533–3539.

[333] Rohrdanz, M. A.; Zheng, W.; Clementi, C. *Annu. Rev. Phys. Chem.* **2013**, *64*, 295–316.

[334] Peters, B. *Annu. Rev. Phys. Chem.* **2016**, *67*, 669–690.

[335] Chen, W.; Tan, A. R.; Ferguson, A. L. *J. Chem. Phys.* **2018**, *149*, 072312.

[336] Ma, C.; Piccinin, S.; Fabris, S. *ACS Catal.* **2012**, *2*, 1500–1506.

[337] Dijkstra, E. W. *Numer. Math* **1959**, *1*, 269–271.

[338] Park, J. M.; Laio, A.; Iannuzzi, M.; Parrinello, M. *J. Am. Chem. Soc.* **2006**, *128*, 11318–11319.

[339] Giorgino, T.; Laio, A.; Rodriguez, A. *Comput. Phys. Commun.* **2017**, *217*, 204–209.

[340] Perdew, J. P.; Ruzsinszky, A.; Csonka, G. I.; Vydrov, O. A.; Scuseria, G. E.; Constantin, L. A.; Zhou, X.; Burke, K. *Phys. Rev. Lett.* **2008**, *100*, 136406.